GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 11, 2004, 11:51:57 ; Search time 19 Seconds
(without alignments)
874.924 Million cell updates/sec

Title: US-10-063-510-6
Perfect score: 1657
Sequence: 1 MARCFSLVLLLTSIWTTRLL.....NPEESKSPSKTTVRCLEAEV 322

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1657	100.0	322	4	US-09-232-160-21	Sequence 21, Appl
2	1657	100.0	322	4	US-09-907-794A-201	Sequence 201, App
3	1657	100.0	322	4	US-09-905-125A-201	Sequence 201, App
4	1657	100.0	322	4	US-09-902-775A-201	Sequence 201, App
5	1651	99.6	322	2	US-08-892-880-2	Sequence 2, Appli
6	1103	66.6	318	4	US-09-724-864-60	Sequence 60, Appl
7	231.5	14.0	363	1	US-07-946-497-7	Sequence 7, Appli
8	231.5	14.0	363	1	US-08-483-322-7	Sequence 7, Appli
9	231.5	14.0	363	2	US-08-478-882-7	Sequence 7, Appli
10	224.5	13.5	339	2	US-08-892-880-3	Sequence 3, Appli
11	222.5	13.4	362	6	5504194-2	Patent No. 5504194
12	222	13.4	503	1	US-07-946-497-2	Sequence 2, Appli
13	222	13.4	503	1	US-08-483-322-2	Sequence 2, Appli
14	222	13.4	503	2	US-08-478-882-2	Sequence 2, Appli
15	207	12.5	361	1	US-07-946-497-6	Sequence 6, Appli
16	207	12.5	361	1	US-08-483-322-6	Sequence 6, Appli
17	207	12.5	361	2	US-08-478-882-6	Sequence 6, Appli
18	161	9.7	90	2	US-08-242-097-3	Sequence 3, Appli
19	161	9.7	90	3	US-09-206-695-3	Sequence 3, Appli
20	161	9.7	90	4	US-09-799-118-3	Sequence 3, Appli
21	145	8.8	277	1	US-08-024-868-2	Sequence 2, Appli
22	145	8.8	277	2	US-08-242-097-2	Sequence 2, Appli
23	145	8.8	277	3	US-09-206-695-2	Sequence 2, Appli
24	145	8.8	277	4	US-09-000-179-1	Sequence 1, Appli
25	145	8.8	277	4	US-09-799-118-2	Sequence 2, Appli
26	145	8.8	277	5	PCT-US96-11995-1	Sequence 1, Appli
27	128.5	7.8	912	5	PCT-US95-03747-2	Sequence 2, Appli
28	121.5	7.3	360	4	US-09-907-794A-213	Sequence 213, App
29	121.5	7.3	360	4	US-09-905-125A-213	Sequence 213, App
30	121.5	7.3	360	4	US-09-902-775A-213	Sequence 213, App
31	121	7.3	528	4	US-09-010-147B-20	Sequence 20, Appl
32	115	6.9	908	5	PCT-US95-03747-3	Sequence 3, Appli
33	114.5	6.9	1257	1	US-08-340-428B-49	Sequence 49, Appl
34	113.5	6.8	97	2	US-08-242-097-5	Sequence 5, Appli
35	113.5	6.8	97	3	US-09-206-695-5	Sequence 5, Appli
36	113.5	6.8	97	4	US-09-799-118-5	Sequence 5, Appli
37	111	6.7	328	1	US-08-225-477B-5	Sequence 5, Appli
38	111	6.7	328	5	PCT-US95-04353-5	Sequence 5, Appli
39	111	6.7	2409	6	5180808-2	Patent No. 5180808
40	107	6.5	371	1	US-08-225-477B-8	Sequence 8, Appli
41	107	6.5	371	5	PCT-US95-04353-8	Sequence 8, Appli
42	106	6.4	329	1	US-08-225-477B-3	Sequence 3, Appli
43	106	6.4	329	5	PCT-US95-04353-3	Sequence 3, Appli
44	104	6.3	98	2	US-08-242-097-4	Sequence 4, Appli
45	104	6.3	98	3	US-09-206-695-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-232-160-21
; Sequence 21, Application US/09232160
; Patent No. 6368794
; GENERAL INFORMATION:
; APPLICANT: Steve Daniel
; APPLICANT: James Gilmore
; APPLICANT: Susan G. Stuart
; APPLICANT: Laura Stuve
; TITLE OF INVENTION: DETECTION OF ALTERED EXPRESSION OF GENES REGULATING CELL
; TITLE OF INVENTION: PROLIFERATION
; FILE REFERENCE: PA-0003 US
; CURRENT APPLICATION NUMBER: US/09/232,160
; CURRENT FILING DATE: 1999-01-15
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PERL Program
; SEQ ID NO 21
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 3044710
US-09-232-160-21

Query Match 100.0%; Score 1657; DB 4; Length 322;
Best Local Similarity 100.0%; Pred. No. 1.4e-159;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy        301 DKNPEESKSPSKTTVRCLEAEV 322
              ||||||||||||||||||||||
```

```
Db        301 DKNPEESKSPSKTTVRCLEAEV 322

RESULT 2
US-09-907-794A-201
; Sequence 201, Application US/09907794A
; Patent No. 6635468
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/907,794A
;  CURRENT FILING DATE:  2001-07-17
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
```

SEQ ID #104

```
; SEQ ID NO 201
;   LENGTH: 322
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic protein
US-09-907-794A-201

  Query Match           100.0%;  Score 1657;  DB 4;  Length 322;
  Best Local Similarity  100.0%;  Pred. No. 1.4e-159;
  Matches  322;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy        301 DKNPEESKSPSKTTVRCLEAEV 322
              ||||||||||||||||||||||
Db        301 DKNPEESKSPSKTTVRCLEAEV 322
```

```
RESULT 3
US-09-905-125A-201
; Sequence 201, Application US/09905125A
; Patent No. 6664376
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/905,125A
;  CURRENT FILING DATE:  2001-07-12
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
```

104

```
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 201
;   LENGTH: 322
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic protein
US-09-905-125A-201

  Query Match            100.0%;  Score 1657;  DB 4;  Length 322;
  Best Local Similarity  100.0%;  Pred. No. 1.4e-159;
  Matches  322;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy        301 DKNPEESKSPSKTTVRCLEAEV 322
              ||||||||||||||||||||||
Db        301 DKNPEESKSPSKTTVRCLEAEV 322


RESULT 4
US-09-902-775A-201
```

```
; Sequence 201, Application US/09902775A
; Patent No. 6686451
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/902,775A
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 201
;   LENGTH: 322
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
```

; OTHER INFORMATION: Synthetic protein
US-09-902-775A-201

Query Match 100.0%; Score 1657; DB 4; Length 322;
Best Local Similarity 100.0%; Pred. No. 1.4e-159;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy        301 DKNPEESKSPSKTTVRCLEAEV 322
              |||||||||||||||||||||||
Db        301 DKNPEESKSPSKTTVRCLEAEV 322
```

RESULT 5
US-08-892-880-2
; Sequence 2, Application US/08892880
; Patent No. 5942417
; GENERAL INFORMATION:
; APPLICANT: NI, JIAN
; APPLICANT: GENTZ, REINER L.
; APPLICANT: DILLON, PATRICK J.
; TITLE OF INVENTION: CD44-LIKE PROTEIN
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/892,880
; FILING DATE: HEREWITH
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/021,762
; FILING DATE: 15-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.0490001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 322 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

; MOLECULE TYPE: protein
US-08-892-880-2

Query Match 99.6%; Score 1651; DB 2; Length 322;
Best Local Similarity 99.7%; Pred. No. 5.7e-159;
Matches 321; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         61 CRLLGLSLAGKDQVETALKASFATCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy        301 DKNPEESKSPSKTTVRCLEAEV 322
              |||||||||||||||||||||||
Db        301 DKNPEESKSPSKTTVRCLEAEV 322
```

RESULT 6
US-09-724-864-60
; Sequence 60, Application US/09724864
; Patent No. 6380362
; GENERAL INFORMATION:
; APPLICANT: Watson, James D
; APPLICANT: Murison, James G.
; TITLE OF INVENTION: Polynucleotides, polypeptides expressed
; TITLE OF INVENTION: by the polynucleotides and methods for their use.
; FILE REFERENCE: 11000.1050U1
; CURRENT APPLICATION NUMBER: US/09/724,864
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: U.S. No. 6380362 60/171,678
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Mouse
US-09-724-864-60

Query Match 66.6%; Score 1103; DB 4; Length 318;
Best Local Similarity 69.7%; Pred. No. 1.9e-103;
Matches 221; Conservative 30; Mismatches 62; Indels 4; Gaps 3;

```
Qy          6 SLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLG 65
              |||||| ||||||  |||:  ::|||  :|||||: || :  | |:||||| |||::||
Db          6 SLVLLLASIWTTRHPVQGADLVQDLSIS-TCRIMGVALVGRNKNPQMNFTEANEACKMLG 64

Qy         66 LSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFA 125
              |:|| :||||:| |: |||||||||||: | || ||  ||:||||| ||||  | |::|
Db         65 LTLASRDQVESAQKSGFETCSYGWVGEQFSVIPRIFSNPRCGKNGKGVLIWNAPSSQKFK 124

Qy        126 AYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTP 185
              |||:|||||| |||||||:||  |: :||  |  ||| || | |  :|| || |   ||
Db        125 AYCHNSSDTWVNSCIPEIVTTFYPVLDTQ--TPATEFSVSSSAYLASSPDSTTPVSATT- 181

Qy        186 PAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTALLVLAL 245
              ||  ||: |: | ||:|||: |  ||:|||| || : ||||||||||||||||||||||
Db        182 RAPPLTSMARKTKKICITEVYTEPITMATETEAFVASGAAFKNEAAGFGGVPTALLVLAL 241
```

```
Qy        246 LFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKTDKNPE 305
              ||||||| |  ||||||||||||| ||||||||||||||||||:| | ||||||| ||||
Db        242 LFFGAAAVLAVCYVKRYVKAFPFTTKNQQKEMIETKVVKEEKADDVNANEESKKTIKNPE 301

Qy        306 ESKSPSKTTVRCLEAEV 322
              |:||| ||||||||||
Db        302 EAKSPPKTTVRCLEAEV 318


RESULT 7
US-07-946-497-7
; Sequence 7, Application US/07946497
; Patent No. 5506119
;  GENERAL INFORMATION:
;    APPLICANT:  HERRLICH, Peter
;    APPLICANT:  PONTA, Helmut
;    APPLICANT:  GUENTHERT, Ursula
;    APPLICANT:  MATZKU, Siegfried
;    APPLICANT:  WENZL, Achim
;    TITLE OF INVENTION:  VARIANT CD44 SURFACE PROTEINS, DNA
;    TITLE OF INVENTION:  SEQUENCES CODING THESE, ANTIBODIES AGAINST THESE PROTEINS,
;    TITLE OF INVENTION:  AS WELL AS THEIR USE IN DIAGNOSIS AND THERAPY
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/946,497
;      FILING DATE:  19921109
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16915/145
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  363 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  mCD44
US-07-946-497-7

  Query Match             14.0%;  Score 231.5;  DB 1;  Length 363;
  Best Local Similarity   24.9%;  Pred. No. 4.5e-15;
  Matches   89;  Conservative   57;  Mismatches  148;  Indels   63;  Gaps   14;

Qy         15 WTTR---LLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGK 71
              | |      |:| ||  ::: : |:||  |:  | |      :: ||| : |:    :|
Db          6 WHTAWGLCLLQLSLAHQQIDLNVTCRYAGVFCVEKNGRYSISRTEAADLCQAFNSTLPTM 65

Qy         72 DQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNS 131
              ||:: ||   |||| ||:: :| ||| || ||  |  |  || |     :  :  ||:|:
Db         66 DQMKLALSKGFETCRYGFI-EGNVVIPRIHPNAICAANHTGVYILVTSNTSHYDTYCFNA 124

Qy        132 SDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDST-YSVASPYST----IPAPT---- 182
              |    |    |:  : |:    |   :  | | ||   | |    | |
Db        125 SAPPEEDC-----TSVTDLPNSFDGPVTITIVNRDGTRYSKKGEYRTHQEDIDASNIIDD 179

Qy        183 ----------TTPPA-PASTSIPRRKKLICVTEVFMETSTMST-ETEPFVENKAAFK--- 227
                       :||       | :|  :      : |   ||::| : :   ::: : :
Db        180 DVSSGSTIEKSTPEGYILHTYLPTEQPTGDQDDSFFIRSTLATRDRDSSKDSRGSSRTVT 239

Qy        228 --NEAAGFGG-----------------VPTALLVLALLFFGAAAGLGFCYVKRYVKAFP 267
                :| ||                   :|  |::|| |     |  |  |     :
Db        240 HGSELAGHSSANQDSGVTTTSGPMRRPQIPEWLIILASL-LALALILAVC-----IAVNS 293

Qy        268 FTNKNQQKEMI---ETKVVKEEKANDSNPNEESKKTDKNPEESKSPSKTTVRCLEAE 321
                    |:|:::       |:: | :: |  | ||  :     :| ||:|  :|: |:
Db        294 RRRCGQKKKLVINGGNGTVEDRKPSELN-GEASKSQEMVHLVNKEPSETPDQCMTAD 349


RESULT 8
US-08-483-322-7
; Sequence 7, Application US/08483322
; Patent No. 5760178
;  GENERAL INFORMATION:
;    APPLICANT:  HERRLICH, Peter
;    APPLICANT:  PONTA, Helmut
;    APPLICANT:  GUENTHERT, Ursula
;    APPLICANT:  MATZKU, Siegfried
;    APPLICANT:  WENZL, Achim
;    TITLE OF INVENTION:  VARIANT CD44 SURFACE PROTEINS, DNA
;    TITLE OF INVENTION:  SEQUENCES CODING THESE, ANTIBODIES AGAINST THESE PROTEINS,
;    TITLE OF INVENTION:  AS WELL AS THEIR USE IN DIAGNOSIS AND THERAPY
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/483,322
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/946,497
;      FILING DATE:  09-NOV-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16915/145
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  363 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  mCD44
US-08-483-322-7

  Query Match             14.0%;  Score 231.5;  DB 1;  Length 363;
  Best Local Similarity   24.9%;  Pred. No. 4.5e-15;
  Matches   89;  Conservative   57;  Mismatches  148;  Indels   63;  Gaps   14;

Qy         15 WTTR---LLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGK 71
              | |      |:| ||  ::: : |:||  |:  | |      :: ||| : |:    :|
Db          6 WHTAWGLCLLQLSLAHQQIDLNVTCRYAGVFCVEKNGRYSISRTEAADLCQAFNSTLPTM 65

Qy         72 DQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNS 131
              ||:: ||   |||| ||:: :| ||| || ||  |  |  || |     :  :  ||:|:
Db         66 DQMKLALSKGFETCRYGFI-EGNVVIPRIHPNAICAANHTGVYILVTSNTSHYDTYCFNA 124
```

```
Qy        132 SDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDST-YSVASPYST----IPAPT---- 182
Db        125 SAPPEEDC-----TSVTDLPNSFDGPVTITIVNRDGTRYSKKGEYRTHQEDIDASNIIDD 179

Qy        183 ----------TTPPA-PASTSIPRRKKLICVTEVFMETSTMST-ETEPFVENKAAFK--- 227
Db        180 DVSSGSTIEKSTPEGYILHTYLPTEQPTGDQDDSFFIRSTLATRDRDSSKDSRGSSRTVT 239

Qy        228 --NEAAGFGG-----------------VPTALLVLALLFFGAAAGLGFCYVKRYVKAFP 267
Db        240 HGSELAGHSSANQDSGVTTTSGPMRRPQIPEWLIILASL-LALALILAVC-----IAVNS 293

Qy        268 FTNKNQQKEMI---ETKVVKEEKANDSNPNEESKKTDKNPEESKSPSKTTVRCLEAE 321
Db        294 RRRCGQKKKLVINGGNGTVEDRKPSELN-GEASKSQEMVHLVNKEPSETPDQCMTAD 349


RESULT 9
US-08-478-882-7
; Sequence 7, Application US/08478882
; Patent No. 5885575
;  GENERAL INFORMATION:
;    APPLICANT:  HERRLICH, Peter
;    APPLICANT:  PONTA, Helmut
;    APPLICANT:  GUENTHERT, Ursula
;    APPLICANT:  MATZKU, Siegfried
;    APPLICANT:  WENZL, Achim
;    TITLE OF INVENTION:  VARIANT CD44 SURFACE PROTEINS, DNA
;    TITLE OF INVENTION:  SEQUENCES CODING THESE, ANTIBODIES AGAINST THESE PROTEINS,
;    TITLE OF INVENTION:  AS WELL AS THEIR USE IN DIAGNOSIS AND THERAPY
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/478,882
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/946,497
;      FILING DATE:  19921109
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16915/145
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  363 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  mCD44
US-08-478-882-7

  Query Match             14.0%;  Score 231.5;  DB 2;  Length 363;
  Best Local Similarity   24.9%;  Pred. No. 4.5e-15;
  Matches   89;  Conservative   57;  Mismatches  148;  Indels   63;  Gaps   14;

Qy         15 WTTR---LLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGK 71
Db          6 WHTAWGLCLLQLSLAHQQIDLNVTCRYAGVFCVEKNGRYSISRTEAADLCQAFNSTLPTM 65

Qy         72 DQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNS 131
Db         66 DQMKLALSKGFETCRYGFI-EGNVVIPRIHPNAICAANHTGVYILVTSNTSHYDTYCFNA 124

Qy        132 SDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDST-YSVASPYST----IPAPT---- 182
Db        125 SAPPEEDC-----TSVTDLPNSFDGPVTITIVNRDGTRYSKKGEYRTHQEDIDASNIIDD 179

Qy        183 ----------TTPPA-PASTSIPRRKKLICVTEVFMETSTMST-ETEPFVENKAAFK--- 227
Db        180 DVSSGSTIEKSTPEGYILHTYLPTEQPTGDQDDSFFIRSTLATRDRDSSKDSRGSSRTVT 239

Qy        228 --NEAAGFGG-----------------VPTALLVLALLFFGAAAGLGFCYVKRYVKAFP 267
Db        240 HGSELAGHSSANQDSGVTTTSGPMRRPQIPEWLIILASL-LALALILAVC-----IAVNS 293

Qy        268 FTNKNQQKEMI---ETKVVKEEKANDSNPNEESKKTDKNPEESKSPSKTTVRCLEAE 321
Db        294 RRRCGQKKKLVINGGNGTVEDRKPSELN-GEASKSQEMVHLVNKEPSETPDQCMTAD 349


RESULT 10
US-08-892-880-3
; Sequence 3, Application US/08892880
; Patent No. 5942417
;  GENERAL INFORMATION:
;    APPLICANT:  NI, JIAN
;    APPLICANT:  GENTZ, REINER L.
;    APPLICANT:  DILLON, PATRICK J.
;    TITLE OF INVENTION:  CD44-LIKE PROTEIN
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
;      STREET:  1100 NEW YORK AVENUE, NW, SUITE 600
;      CITY:  WASHINGTON
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/892,880
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/021,762
;      FILING DATE:  15-JUL-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  STEFFE, ERIC K
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1488.0490001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-371-2600
;      TELEFAX:  202-371-2540
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  339 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-892-880-3

  Query Match             13.5%;  Score 224.5;  DB 2;  Length 339;
  Best Local Similarity   25.1%;  Pred. No. 2.1e-14;
  Matches   83;  Conservative   54;  Mismatches  157;  Indels   37;  Gaps   11;

Qy         15 WTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQV 74
```

```
Db          10 WGLLCLLQLSLAQQQIDLNITCRYAGVFHVEKNGRYSISRTEAADLCEAFNTTLPTMAQM 69

Qy          75 ETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDT 134

Db          70 ELALRKGFETCRYGFI-EGHVVIPRIHPNAICAANNTGVYILLASNTSHYDTYCFNASAP 128

Qy         135 WTNSCIPEIITTKDPIFNTQTATQTTEFIVSDST-YSVASPYST----IPAPT------- 182

Db         129 LEEDC-----TSVTDLPNSFDGPVTITIVNRDGTRYSKKGEYRTHQEDIDASNIIDEDVS 183

Qy         183 -------TTPPA-PASTSIPRRKKLICVTEVFMETSTMST-ETEPFVENKAAFKNEAAGF 233

Db         184 SGSTIEKSTPEGYILHTDLPTSQPTGDRDDAFFIGSTLATGHSSGNQDSGVTTTSGPARR 243

Qy         234 GGVPTALLVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETK---VVKEEKAND 290

Db         244 PQIPEWLIILASL-LALALILAVC-----IAVNSRRRCGQKKKLVINSGNGTVEDRKPSE 297

Qy         291 SNPNEESKKTDKNPEESKSPSKTTVRCLEAE 321

Db         298 LN-GEASKSQEMVHLVNKEPTETPDQFMTAD 327


RESULT 11
5504194-2
;Patent No. 5504194
;    APPLICANT: ST.JOHN, THOMAS P.;GALLATIN, W. MICHAEL;IDZERDA,
;REJEAN
;    TITLE OF INVENTION: LYMPHOCYTE ADHESION RECEPTOR FOR HIGH
;ENDOTHELIUM, CD44
;    NUMBER OF SEQUENCES: 4
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/884,624
;      FILING DATE: 15-MAY-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 628,646
;      FILING DATE: 12-DEC-1990
;      APPLICATION NUMBER: 325,224
;      FILING DATE: 17-MAR-1989
;SEQ ID NO:2:
;      LENGTH: 362
5504194-2

  Query Match            13.4%;  Score 222.5;  DB 6;  Length 362;
  Best Local Similarity   25.5%;  Pred. No. 3.7e-14;
  Matches   92;  Conservative   43;  Mismatches  127;  Indels   99;  Gaps   16;

Qy          20 LVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQVETALK 79

Db          14 LVQLSL--AQIDLNITCRFEGIYHVEKNGRYSISRTEAADLCKAFNSTLPTMAQMEKALS 71

Qy          80 ASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDTWTNSC 139

Db          72 IGFETCRYGFI-EGHVVIPRIHPNSICAANNTGVYILTSNTS-QYDTYCFNASAPPGEDC 129

Qy         140 IPEIITTKDPIFNTQTATQTTEFIVSDST-YSVASPYSTIP------APT---------- 182

Db         130 -----TSVTDLPNAFDGPITITIVNRDGTRYVKKGEYRTNPEDINPSSPTDDDVSSGSSS 184

Qy         183 ----------------TTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAF 226

Db         185 ERSSTLGGYIFYNHFSTSPPIPDEDG---------------PWITDSTDRTPATRDQGAF 229

Qy         227 K----------NEAAGF------GG------------VPTALLVLALLFFGAAAGLGFCY 258

Db         230 DPSGGSHTTHGSESAGHSHGSREGGANTTSGPLRTPQIPEWLIILASL-LALALILAVC- 287

Qy         259 VKRYVKAFPFTNKNQQKEMIETK---VVKEEKANDSNPNEESKKTDK----NPEESKSPS 311

Db         288 ----IAVNSRRRCGQKKKLVINNGNGAVEDRKSSGLN-GEASKSQEMVHLVNKESSETPD 342

Qy         312 K 312

Db         343 Q 343


RESULT 12
US-07-946-497-2
; Sequence 2, Application US/07946497
; Patent No. 5506119
;  GENERAL INFORMATION:
;    APPLICANT:  HERRLICH, Peter
;    APPLICANT:  PONTA, Helmut
;    APPLICANT:  GUENTHERT, Ursula
;    APPLICANT:  MATZKU, Siegfried
;    APPLICANT:  WENZL, Achim
;    TITLE OF INVENTION:  VARIANT CD44 SURFACE PROTEINS, DNA
;    TITLE OF INVENTION:  SEQUENCES CODING THESE, ANTIBODIES AGAINST THESE PROTEINS,
;    TITLE OF INVENTION:  AS WELL AS THEIR USE IN DIAGNOSIS AND THERAPY
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/946,497
;      FILING DATE:  19921109
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16915/145
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  503 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-946-497-2

  Query Match            13.4%;  Score 222;  DB 1;  Length 503;
  Best Local Similarity   25.7%;  Pred. No. 6.7e-14;
  Matches   84;  Conservative   40;  Mismatches  165;  Indels   38;  Gaps   11;

Qy          15 WTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQV 74

Db          10 WGLLCLLQLSLAQQQIDLNITCRYAGVFHVEKNGRYSISRTEAADLCEAFNTTLPTMAQM 69

Qy          75 ETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDT 134

Db          70 ELALRKGFETCRYGFI-EGHVVIPRIHPNAICAANNTGVYILLASNTSHYDTYCFNASAP 128

Qy         135 WTNSCIPEIITTKDPIFNTQTATQTTEFIVSDST-YSVASPYST----IPAPT------- 182

Db         129 LEEDC-----TSVTDLPNSFDGPVTITIVNRDGTRYSKKGEYRTHQEDIDASNIIDEDVS 183

Qy         183 -------TTPPA-PASTSIPRRKKLICVTEVFMETSTMST-ETEPFVENKAAFKNEAAGF 233

Db         184 SGSTIEKSTPEGYILHTDLPTSQPTGDRDDAFFIGSTLATIATTPWVSAHTKQNQERTQW 243

Qy         234 GGV---PTALLVLALLFFGAAAGLGFCYVKRYVK--AFPFTNKNQQKEMIETKVVKEEKA 288

Db         244 NPIHSNPEVLLQTTTRMTDIDRNSTSAHGENWTQEPQPPFNNHEYQDEE-ETPHATSTTW 302

Qy         289 NDSNPNEESKKTDK-----NPEESKSP 310
```

```
Db        303 ADPNSTTEEAATQKEKWFENEWQGKNP 329


RESULT 13
US-08-483-322-2
; Sequence 2, Application US/08483322
; Patent No. 5760178
;  GENERAL INFORMATION:
;    APPLICANT:  HERRLICH, Peter
;    APPLICANT:  PONTA, Helmut
;    APPLICANT:  GUENTHERT, Ursula
;    APPLICANT:  MATZKU, Siegfried
;    APPLICANT:  WENZL, Achim
;    TITLE OF INVENTION:  VARIANT CD44 SURFACE PROTEINS, DNA
;    TITLE OF INVENTION:  SEQUENCES CODING THESE, ANTIBODIES AGAINST THESE PROTEINS,
;    TITLE OF INVENTION:  AS WELL AS THEIR USE IN DIAGNOSIS AND THERAPY
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/483,322
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/946,497
;      FILING DATE:  09-NOV-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16915/145
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  503 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-483-322-2

  Query Match           13.4%;  Score 222;  DB 1;  Length 503;
  Best Local Similarity  25.7%;  Pred. No. 6.7e-14;
  Matches   84;  Conservative   40;  Mismatches  165;  Indels   38;  Gaps   11;

Qy         15 WTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQV 74
Db         10 WGLLCLLQLSLAQQQIDLNITCRYAGVFHVEKNGRYSISRTEAADLCEAFNTTLPTMAQM 69

Qy         75 ETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDT 134
Db         70 ELALRKGFETCRYGFI-EGHVVIPRIHPNAICAANNTGVYILLASNTSHYDTYCFNASAP 128

Qy        135 WTNSCIPEIITTKDPIFNTQTATQTTEFIVSDST-YSVASPYST----IPAPT------- 182
Db        129 LEEDC-----TSVTDLPNSFDGPVTITIVNRDGTRYSKKGEYRTHQEDIDASNIIDEDVS 183

Qy        183 -------TTPPA-PASTSIPRRKKLICVTEVFMETSTMST-ETEPFVENKAAFKNEAAGF 233
Db        184 SGSTIEKSTPEGYILHTDLPTSQPTGDRDDAFFIGSTLATIATTPWVSAHTKQNQERTQW 243

Qy        234 GGV---PTALLVLALLFFGAAAGLGFCYVKRYVK--AFPFTNKNQQKEMIETKVVKEEKA 288

Db        244 NPIHSNPEVLLQTTTRMTDIDRNSTSAHGENWTQEPQPPFNNHEYQDEE-ETPHATSTTW 302

Qy        289 NDSNPNEESKKTDK-----NPEESKSP 310
Db        303 ADPNSTTEEAATQKEKWFENEWQGKNP 329


RESULT 14
US-08-478-882-2
; Sequence 2, Application US/08478882
; Patent No. 5885575
;  GENERAL INFORMATION:
;    APPLICANT:  HERRLICH, Peter
;    APPLICANT:  PONTA, Helmut
;    APPLICANT:  GUENTHERT, Ursula
;    APPLICANT:  MATZKU, Siegfried
;    APPLICANT:  WENZL, Achim
;    TITLE OF INVENTION:  VARIANT CD44 SURFACE PROTEINS, DNA
;    TITLE OF INVENTION:  SEQUENCES CODING THESE, ANTIBODIES AGAINST THESE PROTEINS,
;    TITLE OF INVENTION:  AS WELL AS THEIR USE IN DIAGNOSIS AND THERAPY
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/478,882
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/946,497
;      FILING DATE:  19921109
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16915/145
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  503 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-478-882-2

  Query Match           13.4%;  Score 222;  DB 2;  Length 503;
  Best Local Similarity  25.7%;  Pred. No. 6.7e-14;
  Matches   84;  Conservative   40;  Mismatches  165;  Indels   38;  Gaps   11;

Qy         15 WTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQV 74
Db         10 WGLLCLLQLSLAQQQIDLNITCRYAGVFHVEKNGRYSISRTEAADLCEAFNTTLPTMAQM 69

Qy         75 ETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDT 134
Db         70 ELALRKGFETCRYGFI-EGHVVIPRIHPNAICAANNTGVYILLASNTSHYDTYCFNASAP 128

Qy        135 WTNSCIPEIITTKDPIFNTQTATQTTEFIVSDST-YSVASPYST----IPAPT------- 182
Db        129 LEEDC-----TSVTDLPNSFDGPVTITIVNRDGTRYSKKGEYRTHQEDIDASNIIDEDVS 183

Qy        183 -------TTPPA-PASTSIPRRKKLICVTEVFMETSTMST-ETEPFVENKAAFKNEAAGF 233
```

```
Db        184 SGSTIEKSTPEGYILHTDLPTSQPTGDRDDAFFIGSTLATIATTPWVSAHTKQNQERTQW 243

Qy        234 GGV---PTALLVLALLFFGAAAGLGFCYVKRYVK--AFPFTNKNQQKEMIETKVVKEEKA 288

Db        244 NPIHSNPEVLLQTTTRMTDIDRNSTSAHGENWTQEPQPPFNNHEYQDEE-ETPHATSTTW 302

Qy        289 NDSNPNEESKKTDK-----NPEESKSP 310

Db        303 ADPNSTTEEAATQKEKWFENEWQGKNP 329


RESULT 15
US-07-946-497-6
; Sequence 6, Application US/07946497
; Patent No. 5506119
;  GENERAL INFORMATION:
;    APPLICANT:  HERRLICH, Peter
;    APPLICANT:  PONTA, Helmut
;    APPLICANT:  GUENTHERT, Ursula
;    APPLICANT:  MATZKU, Siegfried
;    APPLICANT:  WENZL, Achim
;    TITLE OF INVENTION:  VARIANT CD44 SURFACE PROTEINS, DNA
;    TITLE OF INVENTION:  SEQUENCES CODING THESE, ANTIBODIES AGAINST THESE PROTEINS,
;    TITLE OF INVENTION:  AS WELL AS THEIR USE IN DIAGNOSIS AND THERAPY
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/946,497
;      FILING DATE:  19921109
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16915/145
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  361 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  hCD44
US-07-946-497-6

  Query Match             12.5%;  Score 207;  DB 1;  Length 361;
  Best Local Similarity   24.4%;  Pred. No. 1.4e-12;
  Matches   83;  Conservative   47;  Mismatches  144;  Indels   66;  Gaps   14;

Qy         25 LRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQVETALKASFET 84

Db         17 LSLAQIDLNITCRFAGVFHVEKNGRYSISRTEAADLCKAFNSTLPTMAQMEKALSIGFET 76

Qy         85 CSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDTWTNSCIPEII 144

Db         77 CRYGFI-EGHVVIPRIHPNSICAANNTGVYILTYNTS-QYDTYCFNASAPPEEDC----- 129

Qy        145 TTKDPIFNTQTATQTTEFIVSDST-YSVASPYSTIPA---PTTTPPAPASTSIPRRKKLI 200

Db        130 TSVTDLPNAFDGPITITIVNRDGTRYVQKGEYRTNPEDIYPSNPTDDDVSSGSSSERSST 189

Qy        201 CVTEVFMETSTM-------------STETEPFVENKAAF----------KNEAAGF---- 233

Db        190 SGGYIFYTFSTVHPIPDEDSPWITDSTDRIPATRDQDTFHPSGGSHTTHESESDGHSHGS 249

Qy        234 --GG------------VPTALLVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIE 279

Db        250 QEGGANTTSGPIRTPQIPEWLIILASL-LALALILAVC-----IAVNSRRRCGQKKKLVI 303

Qy        280 TK---VVKEEKANDSNPNEESKKTDK----NPEESKSPSK 312

Db        304 NSGNGAVEDRKPSGLN-GEASKSQEMVHLVNKESSETPDQ 342


Search completed: August 11, 2004, 11:55:06
Job time : 20 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 11, 2004, 11:54:43 ; Search time 54 Seconds
(without alignments)
1684.820 Million cell updates/sec

Title: US-10-063-510-6
Perfect score: 1657
Sequence: 1 MARCFSLVLLLTSIWTTRLL.....NPEESKSPSKTTVRCLEAEV 322

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 500 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1657	100.0	322	2	AAY13379	Aay13379 Amino aci
2	1657	100.0	322	3	AAY87287	Aay87287 Human sig
3	1657	100.0	322	3	ADC78521	Adc78521 Human PRO
4	1657	100.0	322	4	AAB80247	Aab80247 Human PRO
5	1657	100.0	322	4	AAB87528	Aab87528 Human PRO
6	1657	100.0	322	4	AAB88391	Aab88391 Human mem
7	1657	100.0	322	5	ABG95853	Abg95853 Human sec
8	1657	100.0	322	5	ABB84832	Abb84832 Human PRO
9	1657	100.0	322	5	ABB95438	Abb95438 Human ang
10	1657	100.0	322	6	ABU71625	Abu71625 Human PRO
11	1657	100.0	322	6	ABU71480	Abu71480 Human PRO
12	1657	100.0	322	6	ABU71926	Abu71926 Human sec
13	1657	100.0	322	6	ABO01809	Abo01809 Novel hum
14	1657	100.0	322	6	ABU90878	Abu90878 Novel hum
15	1657	100.0	322	6	ABO33937	Abo33937 Human sec
16	1657	100.0	322	6	ABU71954	Abu71954 Novel hum
17	1657	100.0	322	6	ABU54382	Abu54382 Human sec
18	1657	100.0	322	6	ABO47397	Abo47397 Human sec
19	1657	100.0	322	6	ABU71508	Abu71508 Human sec
20	1657	100.0	322	6	ABU72289	Abu72289 Human PRO
21	1657	100.0	322	6	ABU90962	Abu90962 Human PRO
22	1657	100.0	322	6	ABO27283	Abo27283 Human sec
23	1657	100.0	322	6	ABU64534	Abu64534 Human sec
24	1657	100.0	322	6	ABU67380	Abu67380 Human sec
25	1657	100.0	322	6	ABU92478	Abu92478 Human sec
26	1657	100.0	322	6	ABO14900	Abo14900 Human sec
27	1657	100.0	322	6	ABU81148	Abu81148 Human sec
28	1657	100.0	322	6	ABO53263	Abo53263 Novel hum
29	1657	100.0	322	6	ABU98265	Abu98265 Novel hum
30	1657	100.0	322	6	ABU89270	Abu89270 Novel hum
31	1657	100.0	322	6	ABU82477	Abu82477 Novel hum
32	1657	100.0	322	6	ABU69657	Abu69657 Novel hum
33	1657	100.0	322	6	ABU96441	Abu96441 Human PRO
34	1657	100.0	322	6	ABU72111	Abu72111 Human PRO
35	1657	100.0	322	6	ABO14839	Abo14839 Human sec
36	1657	100.0	322	6	ADB29406	Adb29406 Human sec
37	1657	100.0	322	6	ADB17063	Adb17063 Human tra
38	1657	100.0	322	6	ABO44241	Abo44241 Human sec
39	1657	100.0	322	6	ADA18262	Ada18262 Human sec
40	1657	100.0	322	6	ABO32791	Abo32791 Human sec
41	1657	100.0	322	6	ADA19868	Ada19868 Novel hum
42	1657	100.0	322	6	ADB17251	Adb17251 Human tra
43	1657	100.0	322	6	ABO34851	Abo34851 Human PRO
44	1657	100.0	322	6	ADA16237	Ada16237 Human sec
45	1657	100.0	322	6	ADA20040	Ada20040 Novel hum
46	1657	100.0	322	6	ABO34169	Abo34169 Human sec
47	1657	100.0	322	6	ADA42382	Ada42382 Human sec
48	1657	100.0	322	6	ABO17529	Abo17529 Human PRO
49	1657	100.0	322	6	ADA00337	Ada00337 Human sec
50	1657	100.0	322	7	ADA16661	Ada16661 Human sec
51	1657	100.0	322	7	ADA13090	Ada13090 Human sec
52	1657	100.0	322	7	ADA41958	Ada41958 Human sec
53	1657	100.0	322	7	ADA17305	Ada17305 Human sec
54	1657	100.0	322	7	ADA42808	Ada42808 Human sec
55	1657	100.0	322	7	ABO17590	Abo17590 Human PRO
56	1657	100.0	322	7	ADB85579	Adb85579 Novel hum
57	1657	100.0	322	7	ADB77727	Adb77727 Human sec
58	1657	100.0	322	7	ADB74863	Adb74863 Human sec
59	1657	100.0	322	7	ADB68258	Adb68258 Human PRO
60	1657	100.0	322	7	ADB68065	Adb68065 Human PRO
61	1657	100.0	322	7	ADB90882	Adb90882 Novel hum
62	1657	100.0	322	7	ADC28509	Adc28509 Human sec
63	1657	100.0	322	7	ADC39709	Adc39709 Human sec
64	1657	100.0	322	7	ADC40223	Adc40223 Human sec
65	1657	100.0	322	7	ADC19047	Adc19047 Human sec
66	1657	100.0	322	7	ADC34347	Adc34347 Human sec
67	1657	100.0	322	7	ADC29402	Adc29402 Human sec
68	1657	100.0	322	7	ADC28933	Adc28933 Human sec
69	1657	100.0	322	7	ADC40818	Adc40818 Human sec
70	1657	100.0	322	7	ADC19475	Adc19475 Human sec
71	1657	100.0	322	7	ADC06962	Adc06962 Human PRO
72	1657	100.0	322	7	ADC17141	Adc17141 Mammalian
73	1657	100.0	322	7	ADC33923	Adc33923 Human sec
74	1657	100.0	322	7	ADC12993	Adc12993 Human sec
75	1657	100.0	322	7	ADC14839	Adc14839 Novel hum
76	1657	100.0	322	7	ADC52334	Adc52334 Novel hum
77	1657	100.0	322	7	ADC12445	Adc12445 Human sec
78	1657	100.0	322	7	ADD10321	Add10321 Human sec
79	1657	100.0	322	7	ADD05000	Add05000 Human sec
80	1657	100.0	322	7	ADD11281	Add11281 Human sec
81	1657	100.0	322	7	ADD04006	Add04006 Human sec
82	1657	100.0	322	7	ADD03582	Add03582 Human sec
83	1657	100.0	322	7	ADD37074	Add37074 Human sec
84	1657	100.0	322	7	ADD36010	Add36010 Novel hum
85	1657	100.0	322	7	ADE34834	Ade34834 Human sec
86	1657	100.0	322	8	ADC52144	Adc52144 Novel hum
87	1657	100.0	322	8	ADE79279	Ade79279 Human sec
88	1657	100.0	322	8	ADE79703	Ade79703 Human sec
89	1657	100.0	322	8	ADE73379	Ade73379 Human sec
90	1657	100.0	322	8	ADE41282	Ade41282 Human sec
91	1657	100.0	322	8	ADE71444	Ade71444 Human PDE
92	1657	100.0	322	8	ADE73914	Ade73914 Human sec
93	1654	99.8	344	4	ABB11979	Abb11979 Human PRO
94	1651	99.6	322	2	AAW56249	Aaw56249 Amino aci
95	1651	99.6	322	5	ABB90289	Abb90289 Human pol
96	1511.5	91.2	297	6	ABU69126	Abu69126 Human NOV
97	1218.5	73.5	250	3	AAB34702	Aab34702 Human sec
98	1103	66.6	318	4	AAE05364	Aae05364 Mouse lym

99	821	49.5	255	5	ABB72376	Abb72376	Murine pr
100	592	35.7	116	2	AAY12323	Aay12323	Human 5′
101	277	16.7	58	2	AAY12853	Aay12853	Human 5′
102	261	15.8	69	2	AAY12170	Aay12170	Human 5′
103	229	13.8	364	5	ABB81033	Abb81033	Rat glyco
104	222.5	13.4	362	2	AAR07355	Aar07355	B7 adhesi
105	222.5	13.4	365	6	ABU79109	Abu79109	CD44 prot
106	222	13.4	503	2	AAR14768	Aar14768	Metastasi
107	222	13.4	503	7	ADE57911	Ade57911	Rat Prote
108	208.5	12.6	668	6	ABU04619	Abu04619	Human exp
109	207	12.5	361	2	AAR20816	Aar20816	Haematopo
110	207	12.5	361	2	AAR91444	Aar91444	Human hae
111	207	12.5	361	2	AAW80453	Aaw80453	Human CD4
112	207	12.5	361	2	AAW86200	Aaw86200	Human CD4
113	207	12.5	361	3	AAY96139	Aay96139	Human hae
114	207	12.5	361	4	AAU02448	Aau02448	Human hae
115	207	12.5	361	6	ABU04632	Abu04632	Human exp
116	207	12.5	361	6	ABU04610	Abu04610	Human exp
117	207	12.5	361	6	ABU04638	Abu04638	Human exp
118	207	12.5	361	6	ABU04634	Abu04634	Human exp
119	207	12.5	361	6	ABU04626	Abu04626	Human exp
120	207	12.5	361	6	ABU04630	Abu04630	Human exp
121	207	12.5	361	6	ABU04636	Abu04636	Human exp
122	205	12.4	361	3	AAY99851	Aay99851	Human CD4
123	205	12.4	361	6	ABU04643	Abu04643	Human exp
124	205	12.4	361	6	ABU04609	Abu04609	Human exp
125	205	12.4	361	6	ABU04644	Abu04644	Human exp
126	204	12.3	361	6	ABU04622	Abu04622	Human exp
127	204	12.3	361	6	AAE30338	Aae30338	Human CD4
128	203.5	12.3	608	6	ABP72424	Abp72424	Human CD4
129	203.5	12.3	700	4	AAY97579	Aay97579	Human CD4
130	203.5	12.3	700	6	ABU04640	Abu04640	Human exp
131	203.5	12.3	700	7	ADD90594	Add90594	Human CD4
132	203	12.3	361	7	ADD90592	Add90592	Human CD4
133	203	12.3	436	4	AAY97651	Aay97651	CD44Hextr
134	203	12.3	436	6	ABU04642	Abu04642	Human exp
135	203	12.3	742	6	ABU04653	Abu04653	Human exp
136	203	12.3	742	6	ABU04616	Abu04616	Human exp
137	202.5	12.2	431	4	AAY97650	Aay97650	CD44Hextr
138	202.5	12.2	431	6	ABU04641	Abu04641	Human exp
139	202	12.2	361	6	ABU04607	Abu04607	Human exp
140	202	12.2	675	6	ABU04618	Abu04618	Human exp
141	202	12.2	691	6	ABU04621	Abu04621	Human exp
142	202	12.2	742	6	ABU04620	Abu04620	Human exp
143	202	12.2	742	6	ABU04645	Abu04645	Human exp
144	201	12.1	425	6	ABU04617	Abu04617	Human exp
145	200.5	12.1	699	7	ADD90596	Add90596	Human CD4
146	199.5	12.0	493	2	AAR20817	Aar20817	Epithelia
147	199.5	12.0	493	2	AAR91445	Aar91445	Human epi
148	199.5	12.0	493	2	AAW80454	Aaw80454	Human CD4
149	199.5	12.0	493	2	AAW89151	Aaw89151	Human CD4
150	199.5	12.0	493	3	AAY96140	Aay96140	Human epi
151	199.5	12.0	493	4	AAU02449	Aau02449	Human epi
152	199.5	12.0	493	5	AAU99123	Aau99123	Haematopo
153	199.5	12.0	493	6	ABU04637	Abu04637	Human exp
154	199.5	12.0	493	6	ABU04627	Abu04627	Human exp
155	199.5	12.0	493	6	ABU04639	Abu04639	Human exp
156	199.5	12.0	493	6	ABU04623	Abu04623	Human exp
157	199.5	12.0	493	6	ABU04631	Abu04631	Human exp
158	199.5	12.0	493	6	ABU04633	Abu04633	Human exp
159	199.5	12.0	493	6	ABU04612	Abu04612	Human exp
160	199.5	12.0	493	6	ABU04635	Abu04635	Human exp
161	199.5	12.0	493	6	ABU04613	Abu04613	Human exp
162	199	12.0	338	6	ABU04649	Abu04649	Human exp
163	199	12.0	338	6	ABU04605	Abu04605	Human exp
164	198	11.9	395	6	ABU04615	Abu04615	Human exp
165	197	11.9	719	6	ABU04604	Abu04604	Human exp
166	197	11.9	719	6	ABU04650	Abu04650	Human exp
167	196.5	11.9	699	6	ABU56470	Abu56470	Lung canc
168	196.5	11.9	699	6	ABU04647	Abu04647	Human exp
169	196.5	11.9	699	6	ABU04614	Abu04614	Human exp
170	196.5	11.9	699	6	ABU04608	Abu04608	Human exp
171	196	11.8	293	6	ABU04611	Abu04611	Human exp
172	196	11.8	294	6	ABU04624	Abu04624	Human exp
173	196	11.8	294	6	ABU04646	Abu04646	Human exp
174	194.5	11.7	742	4	ABG17071	Abg17071	Novel hum
175	194	11.7	194	5	AAM48306	Aam48306	Protein R
176	194	11.7	200	5	AAM48307	Aam48307	Protein R
177	194	11.7	273	5	AAM48308	Aam48308	Protein R
178	191.5	11.6	676	6	ABU04602	Abu04602	Human exp
179	191.5	11.6	676	6	ABU04652	Abu04652	Human exp
180	191	11.5	271	6	ABU04648	Abu04648	Human exp
181	191	11.5	271	6	ABU04606	Abu04606	Human exp
182	191	11.5	470	6	ABU04651	Abu04651	Human exp
183	191	11.5	470	6	ABU04603	Abu04603	Human exp
184	188	11.3	170	5	AAM48305	Aam48305	Protein R
185	187	11.3	112	6	ABP73148	Abp73148	Amino aci
186	186.5	11.3	113	6	ABP73150	Abp73150	Amino aci
187	185	11.2	112	6	ABP73151	Abp73151	Amino aci
188	183	11.0	112	6	ABP73152	Abp73152	Amino aci
189	182	11.0	112	6	ABP73149	Abp73149	Amino aci
190	181	10.9	112	6	ABP73153	Abp73153	Amino aci
191	180	10.9	920	4	ABG17067	Abg17067	Novel hum
192	170	10.3	34	4	AAM28056	Aam28056	Peptide #
193	170	10.3	34	5	ABG37308	Abg37308	Human pep
194	150	9.1	510	7	ADB64426	Adb64426	Human pro
195	148.5	9.0	80	6	ABP73154	Abp73154	Deletion
196	145	8.8	277	2	AAR26044	Aar26044	Tumour ne
197	145	8.8	277	2	AAW13654	Aaw13654	Tumour ne
198	145	8.8	277	2	AAW84087	Aaw84087	Tumour ne
199	145	8.8	277	4	AAE02361	Aae02361	Human tum
200	145	8.8	277	5	ABG70870	Abg70870	Human tum
201	145	8.8	277	6	ABR58556	Abr58556	Human can
202	145	8.8	277	6	ABU56653	Abu56653	Lung canc
203	145	8.8	277	7	ADE25781	Ade25781	Human pro
204	134	8.1	897	5	ABJ10588	Abj10588	Human nov
205	134	8.1	2675	5	ABJ10586	Abj10586	Human nov
206	133	8.0	1394	5	AAM47684	Aam47684	Human Hya
207	133	8.0	1416	6	ABG72499	Abg72499	Human 190
208	133	8.0	1431	5	AAM47675	Aam47675	Rat Hyalu
209	133	8.0	1431	6	ABG72498	Abg72498	Rat 175kD
210	133	8.0	1653	6	ABG72514	Abg72514	Human 190
211	131.5	7.9	883	6	ABP72603	Abp72603	Rat mutan
212	131	7.9	649	4	AAB61236	Aab61236	Mature hu
213	131	7.9	649	6	ABO32673	Abo32673	Secreted
214	131	7.9	649	7	ADB90778	Adb90778	Human TAN
215	131	7.9	671	4	AAB61234	Aab61234	Human TAN
216	131	7.9	671	6	ABO32671	Abo32671	Secreted
217	131	7.9	671	7	ADB90776	Adb90776	Human TAN
218	131	7.9	911	4	AAY97583	Aay97583	Human sec
219	131	7.9	911	5	ABG34055	Abg34055	Human Pro
220	131	7.9	911	6	ADA01320	Ada01320	Human PRO
221	131	7.9	911	6	ADA43749	Ada43749	Human sec
222	131	7.9	911	6	ADA43517	Ada43517	Human sec
223	131	7.9	911	6	ADA01192	Ada01192	Human PRO
224	131	7.9	911	7	ADA01076	Ada01076	Human sec
225	131	7.9	911	7	ADA43633	Ada43633	Human sec
226	131	7.9	911	7	ADA06895	Ada06895	Human PRO
227	131	7.9	911	7	ADA08383	Ada08383	Novel hum
228	131	7.9	911	7	ADB99676	Adb99676	Human PRO
229	131	7.9	911	7	ADB86959	Adb86959	Human PRO
230	131	7.9	911	7	ADB66114	Adb66114	Human sec
231	131	7.9	911	7	ADB99792	Adb99792	Human PRO
232	131	7.9	911	7	ADB99447	Adb99447	Novel hum
233	131	7.9	911	7	ADB65998	Adb65998	Human sec
234	131	7.9	911	7	ADC23396	Adc23396	Human tra
235	131	7.9	911	7	ADC26089	Adc26089	Human PRO
236	131	7.9	911	7	ADE62583	Ade62583	Human Pro
237	131	7.9	911	7	ADE04916	Ade04916	Human PRO
238	131	7.9	911	7	ADE11222	Ade11222	Human PRO
239	131	7.9	911	7	ADD88153	Add88153	Human PRO
240	131	7.9	911	7	ADD95448	Add95448	Human sec
241	131	7.9	911	7	ADE06378	Ade06378	Human PRO
242	131	7.9	911	7	ADE38153	Ade38153	Human PRO
243	131	7.9	911	7	ADD88269	Add88269	Human PRO
244	131	7.9	911	7	ADD90850	Add90850	Human sec

245	131	7.9	911	8	ADE51703	Ade51703 Human sec	318	121.5	7.3	360	6	ABU97966	Abu97966 Novel hum
246	131	7.9	911	8	ADE51819	Ade51819 Human sec	319	121.5	7.3	360	6	ABU91672	Abu91672 Novel hum
247	131	7.9	911	8	ADE37677	Ade37677 Human sec	320	121.5	7.3	360	6	ABU71928	Abu71928 Human sec
248	131	7.9	911	8	ADE37561	Ade37561 Human sec	321	121.5	7.3	360	6	ABU89365	Abu89365 Human PRO
249	131	7.9	911	8	ADD95332	Add95332 Human sec	322	121.5	7.3	360	6	ABU86206	Abu86206 Human sec
250	131	7.9	911	8	ADE38032	Ade38032 Human PRO	323	121.5	7.3	360	6	ABU67419	Abu67419 Human sec
251	131	7.9	911	8	ADE76121	Ade76121 Human PRO	324	121.5	7.3	360	6	ABU80447	Abu80447 Human PRO
252	131	7.9	911	8	ADE39444	Ade39444 Human PRO	325	121.5	7.3	360	6	ABO01811	Abo01811 Novel hum
253	131	7.9	911	8	ADE04248	Ade04248 Human PRO	326	121.5	7.3	360	6	ABR99365	Abr99365 Human sec
254	131	7.9	911	8	ADE39845	Ade39845 Human PRO	327	121.5	7.3	360	6	ABR98755	Abr98755 Human sec
255	131	7.9	911	8	ADE19710	Ade19710 Human PRO	328	121.5	7.3	360	6	ABO16278	Abo16278 Human sec
256	131	7.9	911	8	ADE77288	Ade77288 Human sec	329	121.5	7.3	360	6	ABR92178	Abr92178 Human sec
257	131	7.9	911	8	ADE65396	Ade65396 Human PRO	330	121.5	7.3	360	6	ABO18819	Abo18819 Human sec
258	131	7.9	911	8	ADE76005	Ade76005 Human PRO	331	121.5	7.3	360	6	ABR78240	Abr78240 Human sec
259	131	7.9	911	8	ADE37916	Ade37916 Human PRO	332	121.5	7.3	360	6	ABU84976	Abu84976 Novel hum
260	131	7.9	911	8	ADE64526	Ade64526 Human PRO	333	121.5	7.3	360	6	ABO00115	Abo00115 Novel hum
261	131	7.9	911	8	ADE38861	Ade38861 Human PRO	334	121.5	7.3	360	6	ABO11447	Abo11447 Human sec
262	131	7.9	911	8	ADE51935	Ade51935 Human sec	335	121.5	7.3	360	6	ABO02092	Abo02092 Human sec
263	131	7.9	911	8	ADD90966	Add90966 Human sec	336	121.5	7.3	360	6	ABU54384	Abu54384 Human sec
264	131	7.9	911	8	ADE38745	Ade38745 Human PRO	337	121.5	7.3	360	6	ABU88666	Abu88666 Novel hum
265	131	7.9	911	8	ADE37445	Ade37445 Human sec	338	121.5	7.3	360	6	ABU83361	Abu83361 Human sec
266	131	7.9	911	8	ADE06262	Ade06262 Human PRO	339	121.5	7.3	360	6	ABO06162	Abo06162 Novel hum
267	131	7.9	911	8	ADD90121	Add90121 Human sec	340	121.5	7.3	360	6	ABR59198	Abr59198 Human sec
268	131	7.9	911	8	ADE38629	Ade38629 Human PRO	341	121.5	7.3	360	6	ABO09260	Abo09260 Human sec
269	131	7.9	911	8	ADE39560	Ade39560 Human PRO	342	121.5	7.3	360	6	ABO19124	Abo19124 Novel hum
270	131	7.9	911	8	ADD89165	Add89165 Human PRO	343	121.5	7.3	360	6	ABO11142	Abo11142 Human sec
271	131	7.9	911	8	ADD88932	Add88932 Human PRO	344	121.5	7.3	360	6	ABR66760	Abr66760 Human sec
272	131	7.9	911	8	ADE19826	Ade19826 Human PRO	345	121.5	7.3	360	6	ABO15973	Abo15973 Human sec
273	131	7.9	911	8	ADE77404	Ade77404 Human sec	346	121.5	7.3	360	6	ABO13679	Abo13679 Human sec
274	131	7.9	911	8	ADE65280	Ade65280 Human PRO	347	121.5	7.3	360	6	ABO47399	Abo47399 Human sec
275	131	7.9	911	8	ADE39328	Ade39328 Human PRO	348	121.5	7.3	360	6	ABU65582	Abu65582 Human sec
276	131	7.9	911	8	ADE38513	Ade38513 Human sec	349	121.5	7.3	360	6	ABO07430	Abo07430 Human PRO
277	130.5	7.9	883	4	AAB61242	Aab61242 Murine br	350	121.5	7.3	360	6	ABO03617	Abo03617 Human sec
278	130.5	7.9	883	6	ABO32678	Abo32678 Secreted	351	121.5	7.3	360	6	ABR67065	Abr67065 Human sec
279	130.5	7.9	883	6	ABP72604	Abp72604 Rat brain	352	121.5	7.3	360	6	ABO15668	Abo15668 Human sec
280	130.5	7.9	883	7	ADB90787	Adb90787 Mouse bre	353	121.5	7.3	360	6	ABU55949	Abu55949 Human sec
281	130.5	7.9	883	7	ADE62581	Ade62581 Rat Prote	354	121.5	7.3	360	6	ABU65277	Abu65277 Human PRO
282	130.5	7.9	1082	5	ABB90349	Abb90349 Human pol	355	121.5	7.3	360	6	ABU95222	Abu95222 Novel hum
283	129.5	7.8	457	3	AAY93911	Aay93911 A human h	356	121.5	7.3	360	6	ABU71125	Abu71125 Human PRO
284	129.5	7.8	2157	3	AAY93910	Aay93910 A human h	357	121.5	7.3	360	6	ABO07735	Abo07735 Human PRO
285	129.5	7.8	2570	6	ABR82200	Abr82200 Human CLE	358	121.5	7.3	360	6	ABR69976	Abr69976 Human sec
286	128.5	7.8	95	3	AAB08023	Aab08023 The domai	359	121.5	7.3	360	6	ABR69309	Abr69309 Human sec
287	128.5	7.8	95	6	ABU04625	Abu04625 Human exp	360	121.5	7.3	360	6	ABO01450	Abo01450 Human PRO
288	128.5	7.8	912	2	AAR85442	Aar85442 Bovine br	361	121.5	7.3	360	6	ABU81252	Abu81252 Human PRO
289	127.5	7.7	911	6	AAE30340	Aae30340 Human cho	362	121.5	7.3	360	6	ABR60049	Abr60049 Human sec
290	127	7.7	277	6	ABR58557	Abr58557 Human can	363	121.5	7.3	360	6	ABR67784	Abr67784 Human sec
291	127	7.7	277	6	ABU56654	Abu56654 Lung canc	364	121.5	7.3	360	6	ABR65172	Abr65172 Human sec
292	126.5	7.6	457	7	ADB47827	Adb47827 Novel hum	365	121.5	7.3	360	6	ABR68394	Abr68394 Human sec
293	124.5	7.5	482	7	ADB65640	Adb65640 Human pro	366	121.5	7.3	360	6	ABR71806	Abr71806 Human sec
294	121.5	7.3	360	2	AAY13381	Aay13381 Amino aci	367	121.5	7.3	360	6	ABU85286	Abu85286 Human PRO
295	121.5	7.3	360	3	ADC78533	Adc78533 Human PRO	368	121.5	7.3	360	6	ABU88976	Abu88976 Human sec
296	121.5	7.3	360	4	AAB80249	Aab80249 Human PRO	369	121.5	7.3	360	6	ABU83056	Abu83056 Human sec
297	121.5	7.3	360	4	AAU29037	Aau29037 Human PRO	370	121.5	7.3	360	6	ABU94912	Abu94912 Novel hum
298	121.5	7.3	360	4	AAM38965	Aam38965 Human pol	371	121.5	7.3	360	6	ABU90460	Abu90460 Novel hum
299	121.5	7.3	360	6	ABU58413	Abu58413 Human PRO	372	121.5	7.3	360	6	ABU83971	Abu83971 Human sec
300	121.5	7.3	360	6	ABU71627	Abu71627 Human PRO	373	121.5	7.3	360	6	ABU93622	Abu93622 Novel hum
301	121.5	7.3	360	6	ABU87961	Abu87961 Novel hum	374	121.5	7.3	360	6	ABR64867	Abr64867 Human sec
302	121.5	7.3	360	6	ABU84276	Abu84276 Human sec	375	121.5	7.3	360	6	ABR68699	Abr68699 Human sec
303	121.5	7.3	360	6	ABR66150	Abr66150 Human sec	376	121.5	7.3	360	6	ABO06515	Abo06515 Human sec
304	121.5	7.3	360	6	ABR65540	Abr65540 Human sec	377	121.5	7.3	360	6	ABR99060	Abr99060 Human sec
305	121.5	7.3	360	6	ABU99480	Abu99480 Human sec	378	121.5	7.3	360	6	ABU56944	Abu56944 Human PRO
306	121.5	7.3	360	6	ABU82719	Abu82719 Human PRO	379	121.5	7.3	360	6	ABU64536	Abu64536 Human sec
307	121.5	7.3	360	6	ABU89840	Abu89840 Novel hum	380	121.5	7.3	360	6	ABU85896	Abu85896 Novel hum
308	121.5	7.3	360	6	ABU71482	Abu71482 Human PRO	381	121.5	7.3	360	6	ABU67382	Abu67382 Human sec
309	121.5	7.3	360	6	ABR68089	Abr68089 Human sec	382	121.5	7.3	360	6	ABU82183	Abu82183 Novel hum
310	121.5	7.3	360	6	ABU96142	Abu96142 Novel hum	383	121.5	7.3	360	6	ABU87194	Abu87194 Human PRO
311	121.5	7.3	360	6	ABU92573	Abu92573 Human sec	384	121.5	7.3	360	6	ABU83666	Abu83666 Human sec
312	121.5	7.3	360	6	ABO08650	Abo08650 Human sec	385	121.5	7.3	360	6	ABO08040	Abo08040 Human PRO
313	121.5	7.3	360	6	ABO02702	Abo02702 Human sec	386	121.5	7.3	360	6	ABO14902	Abo14902 Human sec
314	121.5	7.3	360	6	ABR74856	Abr74856 Human sec	387	121.5	7.3	360	6	ABU81751	Abu81751 Novel hum
315	121.5	7.3	360	6	ABR94618	Abr94618 Human sec	388	121.5	7.3	360	6	ABU65915	Abu65915 Novel hum
316	121.5	7.3	360	6	ABU85591	Abu85591 Human PRO	389	121.5	7.3	360	6	ABR59744	Abr59744 Human sec
317	121.5	7.3	360	6	ABU98751	Abu98751 Novel hum	390	121.5	7.3	360	6	ABU93932	Abu93932 Novel hum

391	121.5	7.3	360	6	ABU99785	Abu99785 Novel hum
392	121.5	7.3	360	6	ABR66455	Abr66455 Human sec
393	121.5	7.3	360	6	ABR90873	Abr90873 Human sec
394	121.5	7.3	360	6	ABU94300	Abu94300 Human PRO
395	121.5	7.3	360	6	ABU79182	Abu79182 Human PRO
396	121.5	7.3	360	6	ABU86511	Abu86511 Human sec
397	121.5	7.3	360	6	ABU86816	Abu86816 Novel hum
398	121.5	7.3	360	6	ABU94605	Abu94605 Human PRO
399	121.5	7.3	360	6	ABO04532	Abo04532 Human PRO
400	121.5	7.3	360	6	ABR70281	Abr70281 Human sec
401	121.5	7.3	360	6	ABU98446	Abu98446 Human PRO
402	121.5	7.3	360	6	ABR65845	Abr65845 Human sec
403	121.5	7.3	360	6	ABR64562	Abr64562 Human sec
404	121.5	7.3	360	6	ABU79487	Abu79487 Human PRO
405	121.5	7.3	360	6	ABU92878	Abu92878 Human sec
406	121.5	7.3	360	6	ABU95837	Abu95837 Human PRO
407	121.5	7.3	360	6	ABU91057	Abu91057 Novel hum
408	121.5	7.3	360	6	ABU90150	Abu90150 Novel hum
409	121.5	7.3	360	6	ABO09565	Abo09565 Human sec
410	121.5	7.3	360	6	ABO10837	Abo10837 Human sec
411	121.5	7.3	360	6	ABR70891	Abr70891 Human sec
412	121.5	7.3	360	6	ABU87499	Abu87499 Human PRO
413	121.5	7.3	360	6	ABU91367	Abu91367 Human PRO
414	121.5	7.3	360	6	ABU84581	Abu84581 Human sec
415	121.5	7.3	360	6	ABR69671	Abr69671 Human sec
416	121.5	7.3	360	6	ABU80048	Abu80048 Human PRO
417	121.5	7.3	360	6	ABU69659	Abu69659 Novel hum
418	121.5	7.3	360	6	ABU93317	Abu93317 Human PRO
419	121.5	7.3	360	6	ABO09870	Abo09870 Human sec
420	121.5	7.3	360	6	ABO08955	Abo08955 Human sec
421	121.5	7.3	360	6	ABU10523	Abu10523 Human sec
422	121.5	7.3	360	6	ABU95532	Abu95532 Human PRO
423	121.5	7.3	360	6	ABU96741	Abu96741 Novel hum
424	121.5	7.3	360	6	ABR70586	Abr70586 Human sec
425	121.5	7.3	360	6	ABO04937	Abo04937 Novel hum
426	121.5	7.3	360	6	ABO08345	Abo08345 Human sec
427	121.5	7.3	360	6	ABO14841	Abo14841 Human sec
428	121.5	7.3	360	6	ABO05552	Abo05552 Human sec
429	121.5	7.3	360	6	ABR73941	Abr73941 Human sec
430	121.5	7.3	360	6	ABR95533	Abr95533 Human sec
431	121.5	7.3	360	6	ABR80830	Abr80830 Human sec
432	121.5	7.3	360	6	ABR81135	Abr81135 Human sec
433	121.5	7.3	360	6	ABM00831	Abm00831 Human sec
434	121.5	7.3	360	6	ABR88433	Abr88433 Human sec
435	121.5	7.3	360	6	ABM77254	Abm77254 Human sec
436	121.5	7.3	360	6	ABO28738	Abo28738 Human sec
437	121.5	7.3	360	6	ABO31483	Abo31483 Human sec
438	121.5	7.3	360	6	ABM07900	Abm07900 Human sec
439	121.5	7.3	360	6	ABO40380	Abo40380 Human sec
440	121.5	7.3	360	6	ABO35805	Abo35805 Human PRO
441	121.5	7.3	360	6	ABO43944	Abo43944 Human PRO
442	121.5	7.3	360	6	ADA77780	Ada77780 Human sec
443	121.5	7.3	360	6	ABM24739	Abm24739 Human sec
444	121.5	7.3	360	6	ADB29418	Adb29418 Human sec
445	121.5	7.3	360	6	ABO03007	Abo03007 Human sec
446	121.5	7.3	360	6	ABR90263	Abr90263 Human sec
447	121.5	7.3	360	6	ABM17177	Abm17177 Human sec
448	121.5	7.3	360	6	ABR94923	Abr94923 Human sec
449	121.5	7.3	360	6	ABR95228	Abr95228 Human sec
450	121.5	7.3	360	6	ABO21466	Abo21466 Human sec
451	121.5	7.3	360	6	ABR97730	Abr97730 Human sec
452	121.5	7.3	360	6	ABR87518	Abr87518 Human sec
453	121.5	7.3	360	6	ABM77559	Abm77559 Human sec
454	121.5	7.3	360	6	ABM27789	Abm27789 Human sec
455	121.5	7.3	360	6	ABM06070	Abm06070 Human sec
456	121.5	7.3	360	6	ABM03576	Abm03576 Human sec
457	121.5	7.3	360	6	ABM35027	Abm35027 Human sec
458	121.5	7.3	360	6	ABM26264	Abm26264 Human sec
459	121.5	7.3	360	6	ABO48046	Abo48046 Human sec
460	121.5	7.3	360	6	ABR92788	Abr92788 Human sec
461	121.5	7.3	360	6	ABO24549	Abo24549 Human sec
462	121.5	7.3	360	6	ABM11560	Abm11560 Human sec
463	121.5	7.3	360	6	ABM02661	Abm02661 Human sec
464	121.5	7.3	360	6	ABM15957	Abm15957 Human sec
465	121.5	7.3	360	6	ABO27518	Abo27518 Human sec
466	121.5	7.3	360	6	ABM29009	Abm29009 Human sec
467	121.5	7.3	360	6	ABM06985	Abm06985 Human sec
468	121.5	7.3	360	6	ABM21079	Abm21079 Human sec
469	121.5	7.3	360	6	ABM09425	Abm09425 Human sec
470	121.5	7.3	360	6	ABO41295	Abo41295 Human sec
471	121.5	7.3	360	6	ABO36110	Abo36110 Human PRO
472	121.5	7.3	360	6	ABO43639	Abo43639 Human PRO
473	121.5	7.3	360	6	ABM76339	Abm76339 Human sec
474	121.5	7.3	360	6	ABM76035	Abm76035 Human sec
475	121.5	7.3	360	6	ABM25654	Abm25654 Human sec
476	121.5	7.3	360	6	ABM25959	Abm25959 Human sec
477	121.5	7.3	360	6	ABO03312	Abo03312 Human sec
478	121.5	7.3	360	6	ABO02397	Abo02397 Human sec
479	121.5	7.3	360	6	ABR90568	Abr90568 Human sec
480	121.5	7.3	360	6	ABR73636	Abr73636 Human sec
481	121.5	7.3	360	6	ABO16888	Abo16888 Human sec
482	121.5	7.3	360	6	ABR94313	Abr94313 Human sec
483	121.5	7.3	360	6	ABR75820	Abr75820 Human sec
484	121.5	7.3	360	6	ADA18274	Ada18274 Human sec
485	121.5	7.3	360	6	ABO32793	Abo32793 Human sec
486	121.5	7.3	360	6	ABR71196	Abr71196 Human sec
487	121.5	7.3	360	6	ABR93093	Abr93093 Human sec
488	121.5	7.3	360	6	ABR93398	Abr93398 Human sec
489	121.5	7.3	360	6	ABR87823	Abr87823 Human sec
490	121.5	7.3	360	6	ABO27823	Abo27823 Human sec
491	121.5	7.3	360	6	ABO29958	Abo29958 Human sec
492	121.5	7.3	360	6	ABO33167	Abo33167 Human PRO
493	121.5	7.3	360	6	ABM04855	Abm04855 Human sec
494	121.5	7.3	360	6	ABM08815	Abm08815 Human sec
495	121.5	7.3	360	6	ABO36415	Abo36415 Human sec
496	121.5	7.3	360	6	ABO35500	Abo35500 Human PRO
497	121.5	7.3	360	6	ABO39465	Abo39465 Human sec
498	121.5	7.3	360	6	ABM10340	Abm10340 Human sec
499	121.5	7.3	360	6	ABM11865	Abm11865 Human sec
500	121.5	7.3	360	6	ABO52011	Abo52011 Human PRO

ALIGNMENTS

RESULT 1
AAY13379
ID AAY13379 standard; protein; 322 AA.
XX
AC AAY13379;
XX
DT 25-JUN-1999 (first entry)
XX
DE Amino acid sequence of protein PRO263.
XX
KW Secreted protein; transmembrane protein; human; enterocolitis;
KW Zollinger-Ellison syndrome; gastrointestinal ulceration;
KW congenital microvillus atrophy; skin disease; cell growth;
KW abnormal keratinocyte differentiation; psoriasis; epithelial cancer;
KW Parkinson's disease; Alzheimer's disease; ALS; neuropathy; fibromodulin;
KW dermal scarring; Usher Syndrome; Atrophia areata; anti-thrombotic;
KW wound healing; tissue repair.
XX
OS Homo sapiens.
XX
PN WO9914328-A2.
XX
PD 25-MAR-1999.
XX
PF 16-SEP-1998; 98WO-US019330.
XX
PR 17-SEP-1997; 97US-0059113P.
PR 17-SEP-1997; 97US-0059115P.
PR 17-SEP-1997; 97US-0059117P.
PR 17-SEP-1997; 97US-0059119P.
PR 17-SEP-1997; 97US-0059121P.

```
PR   17-SEP-1997;   97US-0059122P.
PR   17-SEP-1997;   97US-0059184P.
PR   18-SEP-1997;   97US-0059263P.
PR   18-SEP-1997;   97US-0059266P.
PR   15-OCT-1997;   97US-0062125P.
PR   17-OCT-1997;   97US-0062285P.
PR   17-OCT-1997;   97US-0062287P.
PR   21-OCT-1997;   97US-0063486P.
PR   24-OCT-1997;   97US-0062814P.
PR   24-OCT-1997;   97US-0062816P.
PR   24-OCT-1997;   97US-0063045P.
PR   24-OCT-1997;   97US-0063120P.
PR   24-OCT-1997;   97US-0063121P.
PR   24-OCT-1997;   97US-0063127P.
PR   24-OCT-1997;   97US-0063128P.
PR   27-OCT-1997;   97US-0063327P.
PR   27-OCT-1997;   97US-0063329P.
PR   28-OCT-1997;   97US-0063541P.
PR   28-OCT-1997;   97US-0063542P.
PR   28-OCT-1997;   97US-0063544P.
PR   28-OCT-1997;   97US-0063549P.
PR   28-OCT-1997;   97US-0063550P.
PR   28-OCT-1997;   97US-0063564P.
PR   29-OCT-1997;   97US-0063435P.
PR   29-OCT-1997;   97US-0063704P.
PR   29-OCT-1997;   97US-0063732P.
PR   29-OCT-1997;   97US-0063734P.
PR   29-OCT-1997;   97US-0063735P.
PR   29-OCT-1997;   97US-0063738P.
PR   29-OCT-1997;   97US-0064215P.
PR   31-OCT-1997;   97US-0063870P.
PR   31-OCT-1997;   97US-0064103P.
PR   03-NOV-1997;   97US-0064248P.
PR   07-NOV-1997;   97US-0064809P.
PR   12-NOV-1997;   97US-0065186P.
PR   17-NOV-1997;   97US-0065846P.
PR   18-NOV-1997;   97US-0065693P.
PR   21-NOV-1997;   97US-0066120P.
PR   21-NOV-1997;   97US-0066364P.
PR   24-NOV-1997;   97US-0066453P.
PR   24-NOV-1997;   97US-0066466P.
PR   24-NOV-1997;   97US-0066511P.
PR   24-NOV-1997;   97US-0066770P.
PR   24-NOV-1997;   97US-0066772P.
PR   25-NOV-1997;   97US-0066840P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Wood WI,  Gurney AL,  Goddard A,  Pennica D,  Chen J,  Yuan J;
XX
DR   WPI; 1999-229533/19.
DR   N-PSDB; AAX52250.
XX
PT   New isolated human genes and polypeptides used in, e.g. treatment of
PT   gastrointestinal ulceration.
XX
PS   Claim 12; Fig 74; 320pp; English.
XX
CC   AAY13344-403 represent secreted and transmembrane human proteins. The
CC   cDNA sequences are obtained from cDNA libraries, prepared from fetal
CC   lung, fetal kidney, fetal brain, fetal liver and fetal retina. The
CC   encoded polypeptides have specific uses based on their homology to known
CC   polypeptides, e.g. PRO211 and PRO217 can be used for disorders associated
CC   with the preservation and maintenance of gastrointestinal mucosa and the
CC   repair of acute and chronic mucosal lesions (e.g. enterocolitis,
CC   Zollinger-Ellison syndrome, gastrointestinal ulceration and congenital
CC   microvillus atrophy), skin diseases associated with abnormal keratinocyte
CC   differentiation (e.g. psoriasis, epithelial cancers such as lung squamous
CC   cell carcinoma of the vulva and gliomas), potent effects on cell growth
CC   and development, diseases related to growth or survival of nerve cells
CC   including Parkinson's disease, Alzheimer's disease, ALS, neuropathies or
CC   cancer. PRO265 can be used as for fibromodulin, e.g. for reducing dermal
CC   scarring. PRO264 can be used as a target for anti-tumor drugs. PRO533 may
CC   be used in the treatment of Usher Syndrome or Atrophia areata; PRO269 can
CC   be used as an anti-thrombotic agent; PRO287 polypeptides and portions may
CC   have therapeutic applications in wound healing and tissue repair; PRO317
CC   can be used for treating problems of the kidney, uterus, endometrium,
CC   blood vessels, or related tissue, e.g. in the heart of genital tract
XX
SQ   Sequence 322 AA;

  Query Match             100.0%;  Score 1657;  DB 2;  Length 322;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-148;
  Matches  322;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy        301 DKNPEESKSPSKTTVRCLEAEV 322
              ||||||||||||||||||||||
Db        301 DKNPEESKSPSKTTVRCLEAEV 322


RESULT 2
AAY87287
ID   AAY87287 standard; protein; 322 AA.
XX
AC   AAY87287;
XX
DT   11-MAY-2000  (first entry)
XX
DE   Human signal peptide containing protein HSPP-64 SEQ ID NO:64.
XX
KW   Human; signal peptide-containing protein; HSPP; diagnosis; cancer;
KW   inflammation; cardiovascular disease; anticancer; anti-inflammatory;
KW   antimicrobial; nootropic; neuroprotective; cardiovascular; hepatotropic;
KW   antiasthmatic; gene therapy; cell proliferation; neurological disorder;
KW   reproductive disorder; developmental disorder; arteriosclerosis;
KW   cirrhosis; psoriasis; acquired immune deficiency syndrome; anaemia;
KW   asthma; Crohn's disease; infection; Alzheimer's disease; schizophrenia;
KW   Parkinson's disease; Huntington's diseases; ovulatory defect;
KW   muscular dystrophy.
XX
OS   Homo sapiens.
XX
PN   WO200000610-A2.
XX
PD   06-JAN-2000.
XX
PF   25-JUN-1999;   99WO-US014484.
XX
PR   26-JUN-1998;   98US-0090762P.
PR   31-JUL-1998;   98US-0094983P.
PR   01-OCT-1998;   98US-0102686P.
PR   11-DEC-1998;   98US-0112129P.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Lal P,  Tang YT,  Gorgone GA,  Corley NC,  Guegler KJ,  Baughn MR;
```

```
PI   Akerblom IE,  Au-Young J,  Yue H,  Patterson C,  Reddy R,  Hillman JL;
PI   Bandman O;
XX
DR   WPI; 2000-160673/14.
DR   N-PSDB; AAZ98172.
XX
PT   New human signal peptide-containing proteins useful in treatment,
PT   prevention and diagnosis of e.g. cancer, inflammation and cardiovascular
PT   disease.
XX
PS   Claim 1; Page 205; 327pp; English.
XX
CC   AAZ98109 to AAZ98242 encode AAY87224 to AAY87357 which represent the
CC   human signal peptide-containing proteins HSPP-1 to HSPP-134. HSPPs have
CC   anticancer, anti-inflammatory, antimicrobial, nootropic, hepatotropic,
CC   neuroprotective, cardiovascular and antiasthmatic activities, and can be
CC   used in gene therapy. HSPPs can be used to treat or prevent disorders
CC   associated with decreased activity or function of HSPP. Antagonists of
CC   HSPP are used to treat or prevent disorders associated with increased
CC   activity or function of HSPP. Such diseases include cell proliferation
CC   (including cancer), inflammation, cardiovascular, neurological,
CC   reproductive or developmental disorders, (e.g. arteriosclerosis,
CC   cirrhosis, psoriasis, acquired immune deficiency syndrome, anaemia,
CC   asthma, Crohn's disease, microbial or other infections, congestive or
CC   ischaemic heart disease, Alzheimer's, Parkinson's or Huntington's
CC   diseases, schizophrenia, ovulatory defects, muscular dystrophy). HSPP
CC   nucleic acids can be used for the recombinant production of HSPP, for
CC   detecting HSPP in standard hybridisation and amplification assays (for
CC   diagnosis and monitoring), in gene therapy, as antisense, triplex-forming
CC   or ribozyme therapeutics, for detecting related sequences or genetic
CC   variations, and for chromosomal mapping. HSPP are also used to raise
CC   specific antibodies (Ab) and to screen for agonists and antagonists
CC   (potential therapeutic agents). Ab are used to diagnose, or monitor, HSPP
CC   -related diseases (in usual immunoassays), as therapeutic antagonists, in
CC   competitive drug screens, and for purification of HSPP from natural
CC   sources
XX
SQ   Sequence 322 AA;

  Query Match             100.0%;  Score 1657;  DB 3;  Length 322;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-148;
  Matches  322;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy        301 DKNPEESKSPSKTTVRCLEAEV 322
              ||||||||||||||||||||||
Db        301 DKNPEESKSPSKTTVRCLEAEV 322


RESULT 3
ADC78521
ID   ADC78521 standard; protein; 322 AA.
XX
AC   ADC78521;
XX
DT   01-JAN-2004  (first entry)
XX
DE   Human PRO263 protein.
XX
KW   antiinflamatory; antiulcer; cytostatic; antipsoriatic; antiparkinsonian;
KW   nootropic; neuroprotective; vasotropic; chemotaxic; angiogenic;
KW   neurotrophic; osteopathic; antiasthmatic; antiarthritic; antirheumatic;
KW   antiarteriosclerotic; cardiant; antidiabetic; cerebroprotective;
KW   thrombolytic; immunomodulator; enterocolitis; Zollinger-Ellison syndrome;
KW   gastrointestinal ulceration; psoriasis; cancer; Parkinson's disease;
KW   Alzheimer's; ALS; neuropathy; dermal scarring; wound healing;
KW   nerve repair; thrombosis; bone; cartilage formation; angiogenesis;
KW   asthma; rheumatoid arthritis; multiple sclerosis; inflammatory disorder;
KW   atherosclerosis; cardiac injury; infertility; premature aging; AIDS;
KW   diabetes; stroke; gene therapy; transgenic; PRO; human.
XX
OS   Homo sapiens.
XX
PN   WO200015796-A2.
XX
PD   23-MAR-2000.
XX
PF   15-SEP-1999;   99WO-US021090.
XX
PR   16-SEP-1998;   98WO-US019330.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Chen J,  Goddard A,  Gurney AL,  Hillan K,  Pennica D,  Wood WI;
PI   Yuan J;
XX
DR   WPI; 2000-271434/23.
DR   N-PSDB; ADC78520.
XX
PT   Novel nucleic acids encoding secreted and transmembrane polypeptides with
PT   homology, e.g. to growth and cancer-associated antigens.
XX
PS   Claim 12; SEQ ID NO 201; 355pp; English.
XX
CC   The invention relates to a novel nucleic acid encoding a PRO polypeptide.
CC   The polypeptides and polynucleotides of the invention may be useful as
CC   research tools and as therapeutics for treating enterocolitis, Zollinger-
CC   Ellison syndrome, gastrointestinal ulceration, psoriasis, cancer,
CC   Parkinson's disease, Alzheimer's disease, ALS, neuropathies, dermal
CC   scarring and wound healing, nerve repair, thrombosis, bone and/or
CC   cartilage formation, angiogenesis, asthma, rheumatoid arthritis, multiple
CC   sclerosis, inflammatory disorders, atherosclerosis, cardiac injury,
CC   infertility, premature aging, AIDS, diabetes complications and stroke.
CC   The molecules may also be utilised during gene therapy procedures and
CC   transgenic animal production. The current sequence is that of the human
CC   PRO protein of the invention.
XX
SQ   Sequence 322 AA;

  Query Match             100.0%;  Score 1657;  DB 3;  Length 322;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-148;
  Matches  322;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy        301 DKNPEESKSPSKTTVRCLEAEV 322
              ||||||||||||||||||||||
Db        301 DKNPEESKSPSKTTVRCLEAEV 322


RESULT 4
AAB80247
ID   AAB80247 standard; protein; 322 AA.
XX
AC   AAB80247;
XX
DT   24-APR-2001  (first entry)
XX
DE   Human PRO263 protein.
XX
KW   Human; PRO; dermatological; antipsoriatic; cytostatic; antiinflammatory;
KW   antiparkinsonian nootropic; neuroprotective; vulnerary; cardiant;
KW   antiangiogenic; vasotropic; antiasthmatic; antirheumatic; cancer;
KW   antiarthritic; antiinfertility; antidiabetic; antiviral; diabetes;
KW   ophthalmological; gene therapy; skin disease; gastrointestinal disorder;
KW   ischaemia; inflammation.
XX
OS   Homo sapiens.
XX
PN   WO200104311-A1.
XX
PD   18-JAN-2001.
XX
PF   22-FEB-2000; 2000WO-US004414.
XX
PR   07-JUL-1999;   99US-0143048P.
PR   26-JUL-1999;   99US-0145698P.
PR   28-JUL-1999;   99US-0146222P.
PR   08-SEP-1999;   99WO-US020594.
PR   13-SEP-1999;   99WO-US020944.
PR   15-SEP-1999;   99WO-US021090.
PR   15-SEP-1999;   99WO-US021547.
PR   05-OCT-1999;   99WO-US023089.
PR   29-NOV-1999;   99WO-US028214.
PR   30-NOV-1999;   99WO-US028313.
PR   02-DEC-1999;   99WO-US028564.
PR   02-DEC-1999;   99WO-US028565.
PR   16-DEC-1999;   99WO-US030095.
PR   20-DEC-1999;   99WO-US030911.
PR   20-DEC-1999;   99WO-US030999.
PR   05-JAN-2000; 2000WO-US000219.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Ashkenazi AJ,  Botstein D,  Desnoyers L,  Eaton DL,  Ferrara N;
PI   Filvaroff E,  Fong S,  Gao W,  Gerber H,  Gerritsen ME,  Goddard A;
PI   Godowski PJ,  Grimaldi CJ,  Gurney AL,  Hillan KJ,  Kljavin IJ;
PI   Mather JP,  Pan J,  Paoni NF,  Roy MA,  Stewart TA,  Tumas D;
PI   Williams PM,  Wood WI;
XX
DR   WPI; 2001-081051/09.
DR   N-PSDB; AAF72408.
XX
PT   Sixty one nucleic acids encoding PRO polypeptides which are useful in the
PT   treatment of skin diseases (e.g. psoriasis), cancers (e.g. lung squamous
PT   cell carcinoma) and neurodegenerative diseases (e.g. Alzheimer's
PT   disease).
XX
PS   Claim 1; Fig 74; 393pp; English.
XX
CC   The present sequence is one of sixty one novel secreted and transmembrane
CC   PRO polypeptides. The PRO polypeptides are useful for treating skin
CC   diseases (e.g. psoriasis), cancers (e.g. lung squamous cell carcinoma),
CC   gastrointestinal disorders (e.g. enterocolitis), neurodegenerative
CC   diseases (e.g. Alzheimer's disease, Parkinson's disease), wound repair,
CC   cardiovascular disorders (e.g. endometrial bleeding angiogenesis,
CC   ischaemias such as coronary ischaemia, atherosclerosis), inflammatory
CC   disorders (e.g. asthma, rheumatoid arthritis, multiple sclerosis),
CC   infertility, AIDS and diabetes and retinal disorders such as retinitis
CC   pigmentosum. The PRO nucleic acids have applications in molecular
CC   biology, including use as hybridization probes, and in chromosome and
CC   gene mapping
XX
SQ   Sequence 322 AA;

  Query Match           100.0%;  Score 1657;  DB 4;  Length 322;
  Best Local Similarity  100.0%;  Pred. No. 6.1e-148;
  Matches  322;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy        301 DKNPEESKSPSKTTVRCLEAEV 322
              ||||||||||||||||||||||
Db        301 DKNPEESKSPSKTTVRCLEAEV 322


RESULT 5
AAB87528
ID   AAB87528 standard; protein; 322 AA.
XX
AC   AAB87528;
XX
DT   15-MAY-2001  (first entry)
XX
DE   Human PRO263.
XX
KW   Human; PRO protein; mapping.
XX
OS   Homo sapiens.
XX
PN   WO200116318-A2.
XX
PD   08-MAR-2001.
XX
PF   24-AUG-2000; 2000WO-US023328.
XX
PR   01-SEP-1999;   99WO-US020111.
PR   15-SEP-1999;   99WO-US021090.
PR   07-DEC-1999;   99US-0169495P.
PR   09-DEC-1999;   99US-0170262P.
PR   11-JAN-2000; 2000US-0175481P.
PR   18-FEB-2000; 2000WO-US004341.
PR   18-FEB-2000; 2000WO-US004342.
PR   22-FEB-2000; 2000WO-US004414.
PR   01-MAR-2000; 2000WO-US005601.
PR   03-MAR-2000; 2000US-0187202P.
PR   21-MAR-2000; 2000US-0191007P.
```

PR 30-MAR-2000; 2000WO-US008439.
PR 25-APR-2000; 2000US-0199397P.
PR 22-MAY-2000; 2000WO-US014042.
PR 05-JUN-2000; 2000US-0209832P.
XX
PA (GETH) GENENTECH INC.
XX
PI Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi CJ, Gurney AL, Watanabe CK, Wood WI;
XX
DR WPI; 2001-183260/18.
DR N-PSDB; AAF92060.
XX
PT Eighty four nucleic acids encoding PRO polypeptides, useful in molecular
PT biology, including use as hybridization probes, and in chromosome and
PT gene mapping.
XX
PS Claim 12; Fig 6; 278pp; English.
XX
CC The present sequence is a human PRO polypeptide (secreted and
CC transmembrane). The PRO protein, and PRO agonists, PRO antagonists or
CC anti-PRO antibodies are useful for preparation of a medicament useful in
CC the treatment of a condition which is responsive to the PRO protein,
CC agonists, antagonists or anti-PRO antibodies. The PRO protein may also be
CC employed as molecular weight markers for protein electrophoresis. The PRO
CC coding sequence has applications in molecular biology, including use as
CC hybridisation probes, and in chromosome and gene mapping
XX
SQ Sequence 322 AA;

Query Match 100.0%; Score 1657; DB 4; Length 322;
Best Local Similarity 100.0%; Pred. No. 6.1e-148;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy        301 DKNPEESKSPSKTTVRCLEAEV 322
              ||||||||||||||||||||||
Db        301 DKNPEESKSPSKTTVRCLEAEV 322
```

RESULT 6
AAB88391
ID AAB88391 standard; protein; 322 AA.
XX
AC AAB88391;
XX
DT 23-MAY-2001 (first entry)
XX
DE Human membrane or secretory protein clone PSEC0135.
XX
KW Human; secretory protein; membrane protein; vaccine; gene therapy;
KW rheumatoid arthritis; diabetes.
XX
OS Homo sapiens.

XX
PN EP1067182-A2.
XX
PD 10-JAN-2001.
XX
PF 07-JUL-2000; 2000EP-00114090.
XX
PR 08-JUL-1999; 99JP-00194179.
PR 11-JAN-2000; 2000JP-00118775.
PR 02-MAY-2000; 2000JP-00183766.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Kawai Y, Sugiyama T, Hayashi K;
XX
DR WPI; 2001-093989/11.
DR N-PSDB; AAF93818.
XX
PT Nucleic acids encoding secretory proteins/membrane proteins, useful in
PT gene therapy or as candidate target molecules in drug development.
XX
PS Claim 1; SEQ ID NO 150; 609pp + Sequence Listing; English.
XX
CC This invention relates to nucleic acid sequences AAF93744 - AAF93916
CC which encode human secretory or membrane proteins represented by AAB88317
CC - AAB88419. Included in the invention are primers AAF93917 - AAF94295 and
CC AAF62232 - AAF62235 which are used to isolate the cDNA sequences of the
CC invention. The invention also includes methods for the production of
CC antibodies directed against the proteins, and cDNA sequences, which can
CC be used in vaccines. The polynucleotide sequences can be used in gene
CC therapy. The polynucleotide sequences and the proteins they encode may be
CC used in the prevention, treatment and diagnosis of diseases associated
CC with inappropriate secretory protein/membrane protein expression. The
CC nucleic acids and complementary sequences may also be used as DNA probes
CC in diagnostic assays (e.g. polymerase chain reactions (PCR)) to detect
CC and quantitate the presence of similar nucleic acid sequences in samples.
CC They may also be used to study the expression and function of secretory
CC proteins/membrane polypeptides and their role in metabolism. The
CC polypeptides may be used as antigens in the production of antibodies
CC against them and in assays to identify modulators (agonists and
CC antagonists) of expression and activity. The antibodies and antagonists
CC may also be used as therapeutic agents to down regulate expression and
CC activity. The antibodies may also be used as diagnostic agents for
CC detecting the presence of the polypeptides in samples (e.g. by enzyme
CC linked immunosorbant assay (ELISA). Examples of diseases which may be
CC treated include rheumatoid arthritis and diabetes
XX
SQ Sequence 322 AA;

Query Match 100.0%; Score 1657; DB 4; Length 322;
Best Local Similarity 100.0%; Pred. No. 6.1e-148;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
```

```
Qy        301 DKNPEESKSPSKTTVRCLEAEV 322
              ||||||||||||||||||||||
Db        301 DKNPEESKSPSKTTVRCLEAEV 322


RESULT 7
ABG95853
ID   ABG95853 standard; protein; 322 AA.
XX
AC   ABG95853;
XX
DT   10-DEC-2002  (first entry)
XX
DE   Human secreted/transmembrane protein PRO263.
XX
KW   Human; secreted protein; transmembrane protein; antirheumatic;
KW   antiarthritic; osteopathic; sports-related joint problem;
KW   articular cartilage defect; osteoarthritis; rheumatoid arthritis.
XX
OS   Homo sapiens.
XX
PN   US2002119130-A1.
XX
PD   29-AUG-2002.
XX
PF   06-DEC-2001; 2001US-00006867.
XX
PR   29-OCT-1997;   97US-0063435P.
PR   29-OCT-1997;   97US-0064215P.
PR   22-APR-1998;   98US-0082797P.
PR   29-APR-1998;   98US-0083495P.
PR   15-MAY-1998;   98US-0085579P.
PR   02-JUN-1998;   98US-0087759P.
PR   04-JUN-1998;   98US-0088021P.
PR   04-JUN-1998;   98US-0088029P.
PR   04-JUN-1998;   98US-0088030P.
PR   10-JUN-1998;   98US-0088734P.
PR   10-JUN-1998;   98US-0088740P.
PR   10-JUN-1998;   98US-0088811P.
PR   10-JUN-1998;   98US-0088824P.
PR   10-JUN-1998;   98US-0088825P.
PR   11-JUN-1998;   98US-0088863P.
PR   12-JUN-1998;   98US-0089105P.
PR   16-JUN-1998;   98US-0089514P.
PR   17-JUN-1998;   98US-0089653P.
PR   19-JUN-1998;   98US-0089952P.
PR   22-JUN-1998;   98US-0090246P.
PR   24-JUN-1998;   98US-0090444P.
PR   25-JUN-1998;   98US-0090688P.
PR   25-JUN-1998;   98US-0090696P.
PR   26-JUN-1998;   98US-0090862P.
PR   02-JUL-1998;   98US-0091628P.
PR   10-AUG-1998;   98US-0096012P.
PR   17-AUG-1998;   98US-0096757P.
PR   18-AUG-1998;   98US-0096949P.
PR   18-AUG-1998;   98US-0096959P.
PR   26-AUG-1998;   98US-0097954P.
PR   26-AUG-1998;   98US-0097971P.
PR   26-AUG-1998;   98US-0097979P.
PR   01-SEP-1998;   98US-0098749P.
PR   10-SEP-1998;   98US-0099741P.
PR   10-SEP-1998;   98US-0099763P.
PR   10-SEP-1998;   98US-0099792P.
PR   10-SEP-1998;   98US-0099812P.
PR   10-SEP-1998;   98US-0099815P.
PR   16-SEP-1998;   98US-0100627P.
PR   16-SEP-1998;   98US-0100662P.
PR   16-SEP-1998;   98WO-US019330.
PR   17-SEP-1998;   98US-0100683P.
PR   17-SEP-1998;   98US-0100684P.
PR   17-SEP-1998;   98US-0100930P.
PR   22-SEP-1998;   98US-0101279P.
PR   23-SEP-1998;   98US-0101475P.
PR   24-SEP-1998;   98US-0101738P.
PR   24-SEP-1998;   98US-0101743P.
PR   24-SEP-1998;   98US-0101916P.
PR   30-SEP-1998;   98US-0102570P.
PR   06-OCT-1998;   98US-0103449P.
PR   08-MAR-1999;   99WO-US005028.
PR   14-MAY-1999;   99WO-US010733.
PR   02-JUN-1999;   99WO-US012252.
PR   01-SEP-1999;   99WO-US020111.
PR   15-SEP-1999;   99WO-US021090.
PR   15-SEP-1999;   99WO-US021194.
PR   22-DEC-1999;   99WO-US030720.
PR   18-FEB-2000; 2000WO-US004341.
PR   18-FEB-2000; 2000WO-US004342.
PR   22-FEB-2000; 2000WO-US004414.
PR   01-MAR-2000; 2000WO-US005601.
PR   30-MAR-2000; 2000WO-US008439.
PR   22-MAY-2000; 2000WO-US014042.
PR   02-JUN-2000; 2000WO-US015264.
PR   23-AUG-2000; 2000WO-US023522.
PR   24-AUG-2000; 2000WO-US023328.
PR   10-NOV-2000; 2000WO-US030873.
PR   01-DEC-2000; 2000WO-US032378.
PR   20-DEC-2000; 2000WO-US034956.
PR   28-FEB-2001; 2001WO-US006520.
PR   01-MAR-2001; 2001WO-US006666.
PR   30-MAY-2001; 2001WO-US017443.
PR   01-JUN-2001; 2001WO-US017800.
PR   20-JUN-2001; 2001WO-US019692.
PR   29-JUN-2001; 2001WO-US021066.
PR   09-JUL-2001; 2001WO-US021735.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Eaton DL,  Filvaroff E,  Gerritsen ME,  Goddard A,  Godowski PJ;
PI   Grimaldi JC,  Gurney AL,  Watanabe CK,  Wood WI;
XX
DR   WPI; 2002-731348/79.
DR   N-PSDB; ABS74380.
XX
PT   New isolated secreted and transmembrane PRO polypeptide useful for
PT   modulating biological activity of a cell, or for treating sports-related
PT   joint problems, osteoarthritis or rheumatoid arthritis.
XX
PS   Claim 20; Fig 6; 399pp; English.
XX
CC   The invention relates to an isolated secreted and transmembrane PRO
CC   polypeptide having 80 % sequence identity to a sequence appearing as
CC   ABG95851-ABG95934 or their associated signal peptide, or a sequence of an
CC   extracellular domain of the proteins with their associated signal peptide
CC   or lacking its associated signal peptide. Also included are the nucleic
CC   acids encoding the proteins, vectors, host cells, fusion proteins and
CC   antibodies which specifically bind to the proteins. The proteins are
CC   useful for detecting a polypeptide designated as A, B, C or D in a sample
CC   suspected of containing an A, B, C or D polypeptide, by contacting the
CC   sample with a polypeptide designated as E, F, G, H or I (or vice versa)
CC   and determining the formation of a A/E, B/F, B/G, C/H or D/I polypeptide
CC   conjugate in the sample, where the formation of the conjugate is
CC   indicative of the presence of an A, B, C or D polypeptide in the sample,
CC   where A is a PRO10272 polypeptide, B is a PRO20110 polypeptide, C is a
CC   PRO10096 polypeptide, D is a PRO19760 polypeptide, E is a PRO5801
CC   polypeptide, F is a PRO1 polypeptide, G is a PRO20040 polypeptide, H is a
CC   PRO20233 polypeptide and I is a PRO1890 polypeptide. The sample comprises
CC   a cell suspected of expressing the A, B, C or D polypeptide. The E, F, G,
CC   H or I polypeptide is labeled with a detectable label or is attached to a
CC   solid support. The proteins are useful for linking a bioactive molecule
CC   to a cell expressing a polypeptide designated as A, B, C or D or E, F, G,
CC   H or I. The bioactive molecule is a toxin, a radiolabel or an antibody.
CC   The bioactive molecule causes death of the cell. A, B, C, D, E, F, G, H,
CC   or I, or antibodies against them are useful for modulating a biological
CC   activity of a cell expressing a polypeptide designated as A, B, C or D or
CC   E, F, G, H, or I. The cell is killed. The proteins are useful for
CC   identifying agonists or antagonists, for the preparation of a medicament
```

```
CC   useful in the treatment of a condition which is responsive to the
CC   proteins, as molecular weight markers for protein electrophoresis
CC   purposes, and as therapeutic agents for treating sports-related joint
CC   problems, articular cartilage defects, osteoarthritis or rheumatoid
CC   arthritis. Nucleic acids encoding the proteins are useful as
CC   hybridisation probes, in chromosome and gene mapping, in the generation
CC   of anti-sense RNA and DNA, for the preparation of the proteins, to
CC   generate transgenic or knockout animals which are useful in the
CC   development and screening of therapeutic useful reagents, for chromosome
CC   identification, and in gene therapy. The antibody is useful as a
CC   therapeutic agent, in a diagnostic assay and for affinity purification of
CC   the protein from recombinant cell culture natural sources. The present
CC   sequence represents a novel secreted or transmembrane protein of the
CC   invention
XX
SQ   Sequence 322 AA;

  Query Match             100.0%;  Score 1657;  DB 5;  Length 322;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-148;
  Matches  322;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy        301 DKNPEESKSPSKTTVRCLEAEV 322
              ||||||||||||||||||||||
Db        301 DKNPEESKSPSKTTVRCLEAEV 322


RESULT 8
ABB84832
ID   ABB84832 standard; protein; 322 AA.
XX
AC   ABB84832;
XX
DT   16-MAY-2002  (first entry)
XX
DE   Human PRO263 protein sequence SEQ ID NO:32.
XX
KW   Human; angiogenesis; cardiant; cytostatic; antiangiogenic; hypotensive;
KW   vulnerary; antiarteriosclerotic; PRO agonist; PRO antagonist; trauma;
KW   gene therapy; cardiovascular disorder; endothelial disorder; cancer;
KW   angiogenic disorder; cardiac hypertrophy; atherosclerosis; hypertension;
KW   age-related macular degeneration; arterial restenosis; angina;
KW   rheumatoid arthritis; myocardial infarction; thrombophlebitis;
KW   lymphangitis; tumour angiogenesis; breast carcinoma; liver carcinoma;
KW   wound healing; chromosome mapping; gene mapping.
XX
OS   Homo sapiens.
XX
PN   WO200200690-A2.
XX
PD   03-JAN-2002.
XX
PF   20-JUN-2001; 2001WO-US019692.
XX
PR   23-JUN-2000; 2000US-0213637P.
PR   20-JUL-2000; 2000US-0219556P.
PR   25-JUL-2000; 2000US-0220624P.
PR   25-JUL-2000; 2000US-0220664P.
PR   28-JUL-2000; 2000WO-US020710.
PR   02-AUG-2000; 2000US-0222695P.
PR   17-AUG-2000; 2000US-00643657.
PR   23-AUG-2000; 2000WO-US023522.
PR   24-AUG-2000; 2000WO-US023328.
PR   07-SEP-2000; 2000US-0230978P.
PR   18-SEP-2000; 2000US-00664610.
PR   18-SEP-2000; 2000US-00665350.
PR   24-OCT-2000; 2000US-0242922P.
PR   08-NOV-2000; 2000US-00709238.
PR   08-NOV-2000; 2000WO-US030952.
PR   10-NOV-2000; 2000WO-US030873.
PR   01-DEC-2000; 2000WO-US032678.
PR   20-DEC-2000; 2000US-00747259.
PR   20-DEC-2000; 2000WO-US034956.
PR   22-JAN-2001; 2001US-00767609.
PR   28-FEB-2001; 2001US-00796498.
PR   28-FEB-2001; 2001WO-US006520.
PR   01-MAR-2001; 2001WO-US006666.
PR   09-MAR-2001; 2001US-00802706.
PR   14-MAR-2001; 2001US-00808689.
PR   22-MAR-2001; 2001US-00816744.
PR   05-APR-2001; 2001US-00828366.
PR   10-MAY-2001; 2001US-00854208.
PR   10-MAY-2001; 2001US-00854280.
PR   25-MAY-2001; 2001US-00866028.
PR   25-MAY-2001; 2001US-00866034.
PR   25-MAY-2001; 2001WO-US017092.
PR   30-MAY-2001; 2001US-00870574.
PR   30-MAY-2001; 2001WO-US017443.
PR   01-JUN-2001; 2001WO-US017800.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Baker KP,  Ferrara N,  Gerber H,  Gerritsen ME,  Goddard A;
PI   Godowski PJ,  Gurney AL,  Hillan KJ,  Marsters SA,  Pan J,  Paoni NF;
PI   Stephan JF,  Watanabe CK,  Williams PM,  Wood WI,  Ye W;
XX
DR   WPI; 2002-090516/12.
DR   N-PSDB; ABL88087.
XX
PT   One hundred and eighty seven nucleic acids encoding PRO polypeptides,
PT   useful in diagnosis and treatment of cardiovascular (e.g. myocardial
PT   infarction), endothelial or angiogenic disorders in a mammal.
XX
PS   Claim 11; Fig 32; 565pp; English.
XX
CC   ABL88072 to ABL88258 encode the PRO proteins given in ABB84817 to
CC   ABB85003. The PRO proteins and polynucleotides have cardiant, cytostatic,
CC   antiangiogenic, hypotensive, vulnerary and antiarteriosclerotic
CC   activities, and can be used in gene therapy. The PRO polynucleotides,
CC   proteins, agonists and antagonists are useful for treating or diagnosing
CC   a cardiovascular, endothelial or angiogenic disorder in a mammal, e.g.
CC   cardiac hypertrophy, trauma, cancer, age-related macular degeneration,
CC   atherosclerosis, hypertension, arterial restenosis, rheumatoid arthritis,
CC   angina, myocardial infarctions, thrombophlebitis, lymphangitis, tumour
CC   angiogenesis (such as breast carcinoma and liver carcinoma) and wound
CC   healing. The PRO polynucleotides have applications in molecular biology,
CC   including use as hybridisation probes, and in chromosome and gene
CC   mapping. ABL88259 to ABL88267 represent primers and probes used in the
CC   exemplification of the present invention
XX
SQ   Sequence 322 AA;

  Query Match             100.0%;  Score 1657;  DB 5;  Length 322;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-148;
  Matches  322;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy        301 DKNPEESKSPSKTTVRCLEAEV 322
                 ||||||||||||||||||||||
Db        301 DKNPEESKSPSKTTVRCLEAEV 322
```

RESULT 9
ABB95438
ID ABB95438 standard; protein; 322 AA.
XX
AC ABB95438;
XX
DT 19-JUL-2002 (first entry)
XX
DE Human angiogenesis related protein PRO263 SEQ ID NO: 32.
XX
KW Human; angiogenesis; PRO protein; cardiovascularisation; wound; cancer;
KW atherosclerosis; cardiac hypertrophy; gene therapy; endothelial disorder;
KW cardiant; cytostatic; antiangiogenic; hypotensive; vulnerary;
KW antiarteriosclerotic.
XX
OS Homo sapiens.
XX
PN WO200208284-A2.
XX
PD 31-JAN-2002.
XX
PF 09-JUL-2001; 2001WO-US021735.
XX
PR 20-JUL-2000; 2000US-0219556P.
PR 25-JUL-2000; 2000US-0220624P.
PR 25-JUL-2000; 2000US-0220664P.
PR 28-JUL-2000; 2000WO-US020710.
PR 02-AUG-2000; 2000US-0222695P.
PR 17-AUG-2000; 2000US-00643657.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 07-SEP-2000; 2000US-0230978P.
PR 18-SEP-2000; 2000US-00664610.
PR 18-SEP-2000; 2000US-00665350.
PR 24-OCT-2000; 2000US-0242922P.
PR 08-NOV-2000; 2000US-00709238.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000WO-US034956.
PR 22-JAN-2001; 2001US-00767609.
PR 28-FEB-2001; 2001US-00796498.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 09-MAR-2001; 2001US-00802706.
PR 14-MAR-2001; 2001US-00808689.
PR 22-MAR-2001; 2001US-00816744.
PR 05-APR-2001; 2001US-00828366.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 25-MAY-2001; 2001US-00866028.
PR 25-MAY-2001; 2001US-00866034.
PR 25-MAY-2001; 2001WO-US017092.
PR 30-MAY-2001; 2001US-00870574.
PR 30-MAY-2001; 2001WO-US017443.
PR 01-JUN-2001; 2001WO-US017800.
PR 20-JUN-2001; 2001WO-US019692.
XX
PA (GETH) GENENTECH INC.
PA (BAKE/) BAKER K P.
PA (FERR/) FERRARA N.
PA (GERB/) GERBER H.
PA (GERR/) GERRITSEN M E.
PA (GODD/) GODDARD A.
PA (GODO/) GODOWSKI P J.
PA (GURN/) GURNEY A L.
PA (HILL/) HILLAN K J.
PA (MARS/) MARSTERS S A.
PA (PANJ/) PAN J.
PA (PAON/) PAONI N F.
PA (STEP/) STEPHAN J F.
PA (WATA/) WATANABE C K.
PA (WILL/) WILLIAMS P M.
PA (WOOD/) WOOD W I.
XX
PI Baker KP, Ferrara N, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Gurney AL, Hillan KJ, Marsters SA, Pan J, Paoni NF;
PI Stephan JF, Watanabe CK, Williams PM, Wood WI, Ye W;
XX
DR WPI; 2002-171999/22.
DR N-PSDB; ABL95576.
XX
PT One hundred and eighty seven nucleic acids encoding PRO polypeptides,
PT useful in diagnosis and treatment of cardiovascular (e.g. myocardial
PT infarction), endothelial or angiogenic disorders in a mammal.
XX
PS Claim 11; Fig 32; 567pp; English.
XX
CC The present invention provides the protein and coding sequences of human
CC PRO proteins. These are useful for treating or diagnosing a
CC cardiovascular, endothelial or angiogenic disorder, including cardiac
CC hypertrophy, trauma, cancer, age-related macular degeneration,
CC atherosclerosis, hypertension, arterial restenosis, rheumatoid arthritis,
CC angina, myocardial infarctions, thrombophlebitis, lymphangitis, tumour
CC angiogenesis (such as breast carcinoma and liver carcinoma) and wound
CC healing. The present sequence is a PRO protein of the invention
XX
SQ Sequence 322 AA;

Query Match 100.0%; Score 1657; DB 5; Length 322;
Best Local Similarity 100.0%; Pred. No. 6.1e-148;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy        301 DKNPEESKSPSKTTVRCLEAEV 322
              |||||||||||||||||||||
Db        301 DKNPEESKSPSKTTVRCLEAEV 322


RESULT 10
ABU71625
ID   ABU71625 standard; protein; 322 AA.
XX
AC   ABU71625;
XX
DT   16-JUN-2003  (first entry)
XX
DE   Human PRO polypeptide #36.
XX
KW   Human; PRO; secreted polypeptide; transmembrane polypeptide;
KW   pathological disorder; cardiac insufficiency disorder; protein secretion;
KW   pancreas; diabetes; gastrointestinal mucosa; mucosal lesion; psoriasis;
KW   skin disease; keratinocyte differentiation; epithelial cancer; tumour;
KW   lung squamous cell carcinoma; epidermoid carcinoma; vulva; glioma;
KW   cytostatic; cardiant; endocrine; antidiabetic; gastrointestinal;
KW   antiulcer; dermatological; vulnerary.
XX
OS   Homo sapiens.
XX
PN   US2002146709-A1.
XX
PD   10-OCT-2002.
XX
PF   18-JUL-2001; 2001US-00909088.
XX
PR   17-SEP-1997;   97US-0059113P.
PR   17-SEP-1997;   97US-0059115P.
PR   17-SEP-1997;   97US-0059117P.
PR   17-SEP-1997;   97US-0059119P.
PR   17-SEP-1997;   97US-0059121P.
PR   17-SEP-1997;   97US-0059122P.
PR   17-SEP-1997;   97US-0059184P.
PR   18-SEP-1997;   97US-0059263P.
PR   18-SEP-1997;   97US-0059266P.
PR   15-OCT-1997;   97US-0062125P.
PR   17-OCT-1997;   97US-0062285P.
PR   17-OCT-1997;   97US-0062287P.
PR   21-OCT-1997;   97US-0063486P.
PR   24-OCT-1997;   97US-0062814P.
PR   24-OCT-1997;   97US-0062816P.
PR   24-OCT-1997;   97US-0063045P.
PR   24-OCT-1997;   97US-0063120P.
PR   24-OCT-1997;   97US-0063121P.
PR   24-OCT-1997;   97US-0063127P.
PR   24-OCT-1997;   97US-0063128P.
PR   27-OCT-1997;   97US-0063327P.
PR   27-OCT-1997;   97US-0063329P.
PR   28-OCT-1997;   97US-0063541P.
PR   28-OCT-1997;   97US-0063542P.
PR   28-OCT-1997;   97US-0063544P.
PR   28-OCT-1997;   97US-0063549P.
PR   28-OCT-1997;   97US-0063550P.
PR   28-OCT-1997;   97US-0063564P.
PR   29-OCT-1997;   97US-0063435P.
PR   29-OCT-1997;   97US-0063704P.
PR   29-OCT-1997;   97US-0063732P.
PR   29-OCT-1997;   97US-0063734P.
PR   29-OCT-1997;   97US-0063735P.
PR   29-OCT-1997;   97US-0063738P.
PR   29-OCT-1997;   97US-0064215P.
PR   31-OCT-1997;   97US-0063870P.
PR   31-OCT-1997;   97US-0064103P.
PR   03-NOV-1997;   97US-0064248P.
PR   07-NOV-1997;   97US-0064809P.
PR   12-NOV-1997;   97US-0065186P.
PR   17-NOV-1997;   97US-0065846P.
PR   18-NOV-1997;   97US-0065693P.
PR   21-NOV-1997;   97US-0066120P.
PR   21-NOV-1997;   97US-0066364P.
PR   24-NOV-1997;   97US-0066453P.
PR   24-NOV-1997;   97US-0066466P.
PR   24-NOV-1997;   97US-0066511P.
PR   24-NOV-1997;   97US-0066770P.
PR   24-NOV-1997;   97US-0066772P.
PR   10-SEP-1998;   98WO-US018824.
PR   14-SEP-1998;   98WO-US019177.
PR   16-SEP-1998;   98WO-US019330.
PR   17-SEP-1998;   98WO-US019437.
PR   01-DEC-1998;   98WO-US025108.
PR   08-SEP-1999;   99WO-US020594.
PR   13-SEP-1999;   99WO-US020944.
PR   15-SEP-1999;   99WO-US021090.
PR   15-SEP-1999;   99WO-US021547.
PR   05-OCT-1999;   99WO-US023089.
PR   29-NOV-1999;   99WO-US028214.
PR   30-NOV-1999;   99WO-US028313.
PR   01-DEC-1999;   99WO-US028301.
PR   02-DEC-1999;   99WO-US028564.
PR   02-DEC-1999;   99WO-US028565.
PR   16-DEC-1999;   99WO-US030095.
PR   20-DEC-1999;   99WO-US030911.
PR   20-DEC-1999;   99WO-US030999.
PR   05-JAN-2000; 2000WO-US000219.
PR   11-FEB-2000; 2000WO-US003565.
PR   22-FEB-2000; 2000WO-US004414.
PR   24-FEB-2000; 2000WO-US005004.
PR   02-MAR-2000; 2000WO-US005841.
PR   20-MAR-2000; 2000WO-US007377.
PR   30-MAR-2000; 2000WO-US008439.
PR   22-MAY-2000; 2000WO-US014042.
PR   02-JUN-2000; 2000WO-US015264.
PR   28-JUL-2000; 2000WO-US020710.
PR   24-AUG-2000; 2000WO-US023328.
PR   18-SEP-2000; 2000US-00665350.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Ashkenazi A,  Botstein D,  Desnoyers L,  Eaton DL,  Ferrara N;
PI   Filvaroff E,  Fong S,  Gao W,  Gerber H,  Gerritsen ME,  Goddard A;
PI   Godowski PJ,  Grimaldi JC,  Gurney AL,  Hillan KJ,  Kljavin IJ;
PI   Mather JP,  Pan J,  Paoni NF,  Roy MA,  Stewart TA,  Tumas D;
PI   Williams PM,  Wood WI;
XX
DR   WPI; 2003-328338/31.
DR   N-PSDB; ACA59060.
XX
PT   Isolated nucleic acid useful for e.g., treating pathological disorders
PT   encodes a secreted or transmembrane protein.
XX
PS   Claim 12; Fig 74; 473pp; English.
XX
CC   The invention relates to human PRO polypeptides (secreted or
CC   transmembrane polypeptides) and the polynucleotides encoding them. The
CC   PRO polypeptides and polynucleotides can be used in treating pathological
CC   disorders and tumours, in therapeutic treatment of cardiac insufficiency
CC   disorders and in therapeutic treatment of disorders involving protein
CC   secretion by the pancreas, including diabetes. They can also be used in
CC   treating disorders associated with the preservation and maintenance of
CC   gastrointestinal mucosa and the repair of acute and chronic mucosal
CC   lesions, and skin diseases associated with abnormal keratinocyte
CC   differentiation (e.g., psoriases, epithelial cancers such as lung
CC   squamous cell carcinoma, epidermoid carcinoma of the vulva and gliomas).
CC   The sequences can be used as molecular markers for protein
CC   electrophoresis purposes and can be utilised in protein-protein binding
CC   assays, biochemical screening assays, immunoassays and cell-based assays.
CC   This sequence represents a human PRO polypeptide of the invention
XX
SQ   Sequence 322 AA;
```

```
  Query Match           100.0%;  Score 1657;  DB 6;  Length 322;
  Best Local Similarity  100.0%;  Pred. No. 6.1e-148;
  Matches  322;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy       61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy      121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy      181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy      241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy      301 DKNPEESKSPSKTTVRCLEAEV 322
            ||||||||||||||||||||||
Db      301 DKNPEESKSPSKTTVRCLEAEV 322
```

RESULT 11
ABU71480
ID ABU71480 standard; protein; 322 AA.
XX
AC ABU71480;
XX
DT 10-JUN-2003 (first entry)
XX
DE Human PRO polypeptide #36.
XX
KW Human; secreted and transmembrane protein; PRO polypeptide; cancer;
KW Alzheimer's disease; ischaemia; cytostatic; nootropic; vasotropic;
KW neuroprotective.
XX
OS Homo sapiens.
XX
PN US2002192659-A1.
XX
PD 19-DEC-2002.
XX
PF 10-JUL-2001; 2001US-00902853.
XX
PR 17-SEP-1997; 97US-0059113P.
PR 17-SEP-1997; 97US-0059115P.
PR 17-SEP-1997; 97US-0059117P.
PR 17-SEP-1997; 97US-0059119P.
PR 17-SEP-1997; 97US-0059121P.
PR 17-SEP-1997; 97US-0059122P.
PR 17-SEP-1997; 97US-0059184P.
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 15-OCT-1997; 97US-0062125P.
PR 17-OCT-1997; 97US-0062285P.
PR 17-OCT-1997; 97US-0062287P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0062814P.
PR 24-OCT-1997; 97US-0062816P.
PR 24-OCT-1997; 97US-0063045P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 24-OCT-1997; 97US-0063127P.
PR 24-OCT-1997; 97US-0063128P.
PR 27-OCT-1997; 97US-0063327P.
PR 27-OCT-1997; 97US-0063329P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063542P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063549P.
PR 28-OCT-1997; 97US-0063550P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063435P.
PR 29-OCT-1997; 97US-0063704P.
PR 29-OCT-1997; 97US-0063732P.
PR 29-OCT-1997; 97US-0063734P.
PR 29-OCT-1997; 97US-0063735P.
PR 29-OCT-1997; 97US-0063738P.
PR 29-OCT-1997; 97US-0064215P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 03-NOV-1997; 97US-0064248P.
PR 07-NOV-1997; 97US-0064809P.
PR 12-NOV-1997; 97US-0065186P.
PR 17-NOV-1997; 97US-0065846P.
PR 18-NOV-1997; 97US-0065693P.
PR 21-NOV-1997; 97US-0066120P.
PR 21-NOV-1997; 97US-0066364P.
PR 24-NOV-1997; 97US-0066453P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066511P.
PR 24-NOV-1997; 97US-0066770P.
PR 24-NOV-1997; 97US-0066772P.
PR 10-SEP-1998; 98WO-US018824.
PR 14-SEP-1998; 98WO-US019177.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98WO-US019437.
PR 01-DEC-1998; 98WO-US025108.
PR 08-SEP-1999; 99WO-US020594.
PR 13-SEP-1999; 99WO-US020944.
PR 15-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021547.
PR 05-OCT-1999; 99WO-US023089.
PR 29-NOV-1999; 99WO-US028214.
PR 30-NOV-1999; 99WO-US028313.
PR 01-DEC-1999; 99WO-US028301.
PR 02-DEC-1999; 99WO-US028564.
PR 02-DEC-1999; 99WO-US028565.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 20-DEC-1999; 99WO-US030999.
PR 05-JAN-2000; 2000WO-US000219.
PR 11-FEB-2000; 2000WO-US003565.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US005004.
PR 02-MAR-2000; 2000WO-US005841.
PR 20-MAR-2000; 2000WO-US007377.
PR 30-MAR-2000; 2000WO-US008439.
PR 22-MAY-2000; 2000WO-US014042.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 24-AUG-2000; 2000WO-US023328.
PR 18-SEP-2000; 2000US-00665350.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi A, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
PI Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ, Kljavin IJ;
PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
PI Williams PM, Wood WI;
XX
DR WPI; 2003-361832/34.
DR N-PSDB; ACA58457.
XX
PT New isolated nucleic acid encoding a PRO polypeptide, e.g. PRO245 or
PT PRO1868, useful in molecular biology, chromosome and gene mapping, in
PT generating antisense RNA and DNA, and in gene therapy.

```
XX
PS   Claim 12; Fig 74; 474pp; English.
XX
CC   The present invention relates to the isolation of novel human secreted
CC   and transmembrane proteins (PRO polypeptides), and the polynucleotide
CC   sequences encoding them. The polynucleotide sequences are useful in
CC   molecular biology, as hybridisation probes, in chromosome and gene
CC   mapping, in generating antisense RNA and DNA, and in gene therapy. The
CC   polynucleotide sequences may also be used in preparing PRO polypeptides
CC   by recombinant techniques, and in generating either transgenic animals or
CC   knock-out animals which, in turn, are useful in the development and
CC   screening of therapeutically useful reagents. The PRO polypeptides or
CC   their antibodies are useful in preparing a medicament for treating a
CC   condition responsive to the polypeptide or antibody, such as cancer,
CC   Alzheimer's disease or ischaemia, and in various diagnostic assays.
CC   ABU71445-ABU71505 represent human PRO polypeptides of the invention
XX
SQ   Sequence 322 AA;

  Query Match             100.0%;  Score 1657;  DB 6;  Length 322;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-148;
  Matches  322;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy        301 DKNPEESKSPSKTTVRCLEAEV 322
              ||||||||||||||||||||||
Db        301 DKNPEESKSPSKTTVRCLEAEV 322


RESULT 12
ABU71926
ID   ABU71926 standard; protein; 322 AA.
XX
AC   ABU71926;
XX
DT   12-JUN-2003  (first entry)
XX
DE   Human secreted/transmembrane protein PRO263.
XX
KW   Human; secreted protein; transmembrane protein; PRO; gene therapy;
KW   chromosome identification; chromosome marker.
XX
OS   Homo sapiens.
XX
PN   US2003003530-A1.
XX
PD   02-JAN-2003.
XX
PF   11-JUL-2001; 2001US-00904011.
XX
PR   17-SEP-1997;   97US-0059113P.
PR   17-SEP-1997;   97US-0059115P.
PR   17-SEP-1997;   97US-0059117P.
PR   17-SEP-1997;   97US-0059119P.
PR   17-SEP-1997;   97US-0059121P.
PR   17-SEP-1997;   97US-0059122P.
PR   17-SEP-1997;   97US-0059184P.
PR   18-SEP-1997;   97US-0059263P.
PR   18-SEP-1997;   97US-0059266P.
PR   15-OCT-1997;   97US-0062125P.
PR   17-OCT-1997;   97US-0062285P.
PR   17-OCT-1997;   97US-0062287P.
PR   21-OCT-1997;   97US-0063486P.
PR   24-OCT-1997;   97US-0062814P.
PR   24-OCT-1997;   97US-0062816P.
PR   24-OCT-1997;   97US-0063045P.
PR   24-OCT-1997;   97US-0063120P.
PR   24-OCT-1997;   97US-0063121P.
PR   24-OCT-1997;   97US-0063127P.
PR   24-OCT-1997;   97US-0063128P.
PR   27-OCT-1997;   97US-0063327P.
PR   27-OCT-1997;   97US-0063329P.
PR   28-OCT-1997;   97US-0063541P.
PR   28-OCT-1997;   97US-0063542P.
PR   28-OCT-1997;   97US-0063544P.
PR   28-OCT-1997;   97US-0063549P.
PR   28-OCT-1997;   97US-0063550P.
PR   28-OCT-1997;   97US-0063564P.
PR   29-OCT-1997;   97US-0063435P.
PR   29-OCT-1997;   97US-0063704P.
PR   29-OCT-1997;   97US-0063732P.
PR   29-OCT-1997;   97US-0063734P.
PR   29-OCT-1997;   97US-0063735P.
PR   29-OCT-1997;   97US-0063738P.
PR   29-OCT-1997;   97US-0064215P.
PR   31-OCT-1997;   97US-0063870P.
PR   31-OCT-1997;   97US-0064103P.
PR   03-NOV-1997;   97US-0064248P.
PR   07-NOV-1997;   97US-0064809P.
PR   12-NOV-1997;   97US-0065186P.
PR   17-NOV-1997;   97US-0065846P.
PR   18-NOV-1997;   97US-0065693P.
PR   21-NOV-1997;   97US-0066120P.
PR   21-NOV-1997;   97US-0066364P.
PR   24-NOV-1997;   97US-0066453P.
PR   24-NOV-1997;   97US-0066466P.
PR   24-NOV-1997;   97US-0066511P.
PR   24-NOV-1997;   97US-0066770P.
PR   24-NOV-1997;   97US-0066772P.
PR   10-SEP-1998;   98WO-US018824.
PR   14-SEP-1998;   98WO-US019177.
PR   16-SEP-1998;   98WO-US019330.
PR   17-SEP-1998;   98WO-US019437.
PR   01-DEC-1998;   98WO-US025108.
PR   08-SEP-1999;   99WO-US020594.
PR   13-SEP-1999;   99WO-US020944.
PR   15-SEP-1999;   99WO-US021090.
PR   15-SEP-1999;   99WO-US021547.
PR   05-OCT-1999;   99WO-US023089.
PR   29-NOV-1999;   99WO-US028214.
PR   30-NOV-1999;   99WO-US028313.
PR   01-DEC-1999;   99WO-US028301.
PR   02-DEC-1999;   99WO-US028564.
PR   02-DEC-1999;   99WO-US028565.
PR   16-DEC-1999;   99WO-US030095.
PR   20-DEC-1999;   99WO-US030911.
PR   20-DEC-1999;   99WO-US030999.
PR   05-JAN-2000; 2000WO-US000219.
PR   11-FEB-2000; 2000WO-US003565.
PR   22-FEB-2000; 2000WO-US004414.
PR   24-FEB-2000; 2000WO-US005004.
PR   02-MAR-2000; 2000WO-US005841.
PR   20-MAR-2000; 2000WO-US007377.
PR   30-MAR-2000; 2000WO-US008439.
PR   22-MAY-2000; 2000WO-US014042.
PR   02-JUN-2000; 2000WO-US015264.
PR   28-JUL-2000; 2000WO-US020710.
```

PR 24-AUG-2000; 2000WO-US023328.
PR 18-SEP-2000; 2000US-00665350.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi A, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
PI Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ, Kljavin IJ;
PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
PI Williams PM, Wood WI;
XX
DR WPI; 2003-329602/31.
DR N-PSDB; ACA60164.
XX
PT New transmembrane polypeptides and nucleic acids encoding the
PT polypeptides, useful in gene therapy, in chromosome identification, as
PT chromosome markers, in generating probes and in tissue typing.
XX
PS Claim 12; Fig 74; 484pp; English.
XX
CC The invention relates to an isolated nucleic acid with at least 80%
CC nucleic acid sequence identity to a nucleotide sequence encoding one of
CC 61 secreted/transmembrane polypeptides, or PRO polypeptides or encoding a
CC PRO protein extracellular domain. Also included are a vector comprising
CC the PRO nucleic acid, a host cell comprising the vector, producing a PRO
CC polypeptide (by culturing the host cell for the expression of the PRO
CC polypeptide, and recovering the PRO polypeptide from the cell culture),
CC an isolated PRO polypeptide (having at least 80% sequence identity to: (
CC a) an amino acid sequence selected from the 61 PRO proteins; (b) an amino
CC acid sequence encoded by a nucleic acid molecule deposited with an ATCC
CC number (detailed in the specification); or (c) an extracellular domain of
CC a PRO polypeptide or to a PRO polypeptide lacking its associated signal
CC peptide), a chimaeric molecule comprising a PRO polypeptide of fused to a
CC heterologous amino acid sequence, an anti-PRO antibody, detecting a
CC PRO245 or PRO1868 in a sample suspected of containing the polypeptide,
CC linking a bioactive molecule to a cell expressing a PRO245 or PRO1868 and
CC modulating at least one biological activity of a cell expressing a PRC245
CC or PRO1868. Nucleic acids which encode PRO can be used to generate either
CC transgenic animals or knock-out animals which may be used in the
CC development and screening of therapeutically useful reagents. The nucleic
CC acids may also be used in gene therapy, in chromosome identification, as
CC chromosome markers, or in generating probes. The PRO polypeptides are
CC useful as molecular markers for protein electrophoresis, and the isolated
CC nucleic acids may be used for recombinantly expressing those markers. The
CC PRO polypeptides and nucleic acids may also be used in tissue typing.
CC Anti-PRO antibodies are useful in diagnostic assays for PRO, and in
CC affinity purification of PRO from recombinant cell culture or natural
CC sources. The present sequence represents a PRO protein
XX
SQ Sequence 322 AA;

Query Match 100.0%; Score 1657; DB 6; Length 322;
Best Local Similarity 100.0%; Pred. No. 6.1e-148;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy        301 DKNPEESKSPSKTTVRCLEAEV 322
              ||||||||||||||||||||||
Db        301 DKNPEESKSPSKTTVRCLEAEV 322
```

RESULT 13
ABO01809
ID ABO01809 standard; protein; 322 AA.
XX
AC ABO01809;
XX
DT 07-AUG-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO263.
XX
KW Human; secreted and transmembrane protein; PRO; pharmaceutical;
KW diagnostic; biosensor; bioreactor; Parkinson's disease;
KW Alzheimer's disease; inflammation; nephritis; wound healing;
KW nerve repair; collateral blood vessel formation; cancer;
KW colorectal cancer; haemorrhage; rheumatoid arthritis; diabetes;
KW cirrhosis; fibrosis; restenosis; dermal fibrotic condition; keloid;
KW scarring; ischaemia; stroke; hypertension; heart attack; atherosclerosis;
KW infertility; gene therapy.
XX
OS Homo sapiens.
XX
PN US2002197671-A1.
XX
PD 26-DEC-2002.
XX
PF 17-JUL-2001; 2001US-00907824.
XX
PR 17-SEP-1997; 97US-0059113P.
PR 17-SEP-1997; 97US-0059115P.
PR 17-SEP-1997; 97US-0059117P.
PR 17-SEP-1997; 97US-0059119P.
PR 17-SEP-1997; 97US-0059121P.
PR 17-SEP-1997; 97US-0059122P.
PR 17-SEP-1997; 97US-0059184P.
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 15-OCT-1997; 97US-0062125P.
PR 17-OCT-1997; 97US-0062285P.
PR 17-OCT-1997; 97US-0062287P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0062814P.
PR 24-OCT-1997; 97US-0062816P.
PR 24-OCT-1997; 97US-0063045P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 24-OCT-1997; 97US-0063127P.
PR 24-OCT-1997; 97US-0063128P.
PR 27-OCT-1997; 97US-0063327P.
PR 27-OCT-1997; 97US-0063329P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063542P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063549P.
PR 28-OCT-1997; 97US-0063550P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063435P.
PR 29-OCT-1997; 97US-0063704P.
PR 29-OCT-1997; 97US-0063732P.
PR 29-OCT-1997; 97US-0063734P.
PR 29-OCT-1997; 97US-0063735P.
PR 29-OCT-1997; 97US-0063738P.
PR 29-OCT-1997; 97US-0064215P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 03-NOV-1997; 97US-0064248P.
PR 07-NOV-1997; 97US-0064809P.

PR 12-NOV-1997; 97US-0065186P.
PR 17-NOV-1997; 97US-0065846P.
PR 18-NOV-1997; 97US-0065693P.
PR 21-NOV-1997; 97US-0066120P.
PR 21-NOV-1997; 97US-0066364P.
PR 24-NOV-1997; 97US-0066453P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066511P.
PR 24-NOV-1997; 97US-0066770P.
PR 24-NOV-1997; 97US-0066772P.
PR 10-SEP-1998; 98WO-US018824.
PR 14-SEP-1998; 98WO-US019177.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98WO-US019437.
PR 01-DEC-1998; 98WO-US025108.
PR 08-SEP-1999; 99WO-US020594.
PR 13-SEP-1999; 99WO-US020944.
PR 15-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021547.
PR 05-OCT-1999; 99WO-US023089.
PR 29-NOV-1999; 99WO-US028214.
PR 30-NOV-1999; 99WO-US028313.
PR 01-DEC-1999; 99WO-US028301.
PR 02-DEC-1999; 99WO-US028564.
PR 02-DEC-1999; 99WO-US028565.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 20-DEC-1999; 99WO-US030999.
PR 05-JAN-2000; 2000WO-US000219.
PR 11-FEB-2000; 2000WO-US003565.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US005004.
PR 02-MAR-2000; 2000WO-US005841.
PR 20-MAR-2000; 2000WO-US007377.
PR 30-MAR-2000; 2000WO-US008439.
PR 22-MAY-2000; 2000WO-US014042.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 24-AUG-2000; 2000WO-US023328.
PR 18-SEP-2000; 2000US-00665350.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi A, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
PI Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ, Kljavin IJ;
PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
PI Williams PM, Wood WI;
XX
DR WPI; 2003-370793/35.
DR N-PSDB; ACD07564.
XX
PT New genes and secreted and transmembrane polypeptides (e.g. PRO245 or
PT PRO335), useful for treating or diagnosing e.g. Alzheimer's disease,
PT cancers, hemorrhage, rheumatoid arthritis, diabetes, cirrhosis, ischemia
PT or strokes.
XX
PS Claim 12; Fig 74; 482pp; English.
XX
CC The invention describes a new isolated nucleic acid molecule comprising
CC the full length coding sequence of the DNA deposited with the American
CC Type Culture Collection (e.g. ATCC Deposit No. 209258) ,or a sequence
CC with at least 80% identity to a DNA encoding a PRO polypeptide comprising
CC any of 61 sequences having 164-1119 amino acids fully defined in the
CC specification. The PRO polypeptides or polynucleotides are useful as
CC pharmaceuticals, diagnostics, biosensors or bioreactors. These are
CC particularly useful for detecting or treating e.g. Parkinson's disease,
CC Alzheimer's disease, inflammations, nephritis, wound healing, nerve
CC repair, collateral blood vessel formation, cancers (e.g. colorectal
CC cancer), haemorrhage (or reduce risk for haemorrhage), rheumatoid
CC arthritis, diabetes, cirrhosis of the liver, fibrosis of the lungs,
CC restenosis, dermal fibrotic conditions (e.g. keloids or scarring),
CC ischaemia, strokes, hypertension, heart attacks, atherosclerosis, or
CC infertility in mammals (e.g. humans, dogs, cats, cattle, horses, sheep,
CC pigs, goats, or rabbits) The PRO polypeptides are useful as targets for
CC therapeutic intervention in these diseases, and diagnostic determination
CC of the presence of these diseases. The PRO polypeptides are also useful
CC as molecular weight markers, or for chromosome identification. The PRO
CC genes are useful as hybridisation probes, or for screening libraries of
CC human cDNA, genomic DNA or mRNA. The PRO genes may also be used in gene
CC therapy, particularly for replacing a defective gene. This is the amino
CC acid sequence of a novel human secreted and transmembrane PRO polypeptide
XX
SQ Sequence 322 AA;

```
  Query Match           100.0%;  Score 1657;  DB 6;  Length 322;
  Best Local Similarity  100.0%;  Pred. No. 6.1e-148;
  Matches  322;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy        301 DKNPEESKSPSKTTVRCLEAEV 322
              ||||||||||||||||||||||
Db        301 DKNPEESKSPSKTTVRCLEAEV 322
```

RESULT 14
ABU90878
ID ABU90878 standard; protein; 322 AA.
XX
AC ABU90878;
XX
DT 11-JUL-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO263.
XX
KW Human; secreted and transmembrane protein; PRO; antibody therapy;
KW pharmaceutical; diagnostic; biosensor; bioreactor.
XX
OS Homo sapiens.
XX
PN US2003018173-A1.
XX
PD 23-JAN-2003.
XX
PF 01-MAY-2002; 2002US-00063515.
XX
PR 06-DEC-2001; 2001US-00006867.
XX
PA (GETH) GENENTECH INC.
XX
PI Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Watanabe CK, Wood WI;
XX
DR WPI; 2003-401702/38.
DR N-PSDB; ACA91166.
XX
PT New antibody useful for identifying PRO polypeptides, for affinity

PT purification of PRO polypeptides, and for preparing a medicament for
PT diagnosing or treating conditions responsive to the antibody or PRO
PT polypeptide.
XX
PS Disclosure; Fig 6; 345pp; English.
XX
CC The invention describes an antibody that specifically binds to a PRO
CC polypeptide having a fully defined amino acid sequence given in the
CC specification. The antibody is useful in identifying PRO polypeptides
CC useful for various industrial applications, including pharmaceuticals,
CC diagnostics, biosensors and bioreactors. The antibody is also used for
CC affinity purification of PRO polypeptides from recombinant cell culture
CC or natural sources. The antibody, PRO polypeptide, or its agonists or
CC antagonists, may be used for preparing a medicament for diagnosing or
CC treating a condition responsive to the antibody, PRO polypeptide, or its
CC agonists or antagonists. This is the amino acid sequence of a novel human
CC secreted and transmembrane PRO polypeptide
XX
SQ Sequence 322 AA;

Query Match 100.0%; Score 1657; DB 6; Length 322;
Best Local Similarity 100.0%; Pred. No. 6.1e-148;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
Db 1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy 61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
Db 61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy 121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
Db 121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy 181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
Db 181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy 241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
Db 241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy 301 DKNPEESKSPSKTTVRCLEAEV 322
Db 301 DKNPEESKSPSKTTVRCLEAEV 322

RESULT 15
ABO33937
ID ABO33937 standard; protein; 322 AA.
XX
AC ABO33937;
XX
DT 18-SEP-2003 (first entry)
XX
DE Human secreted/transmembrane protein PRO263.
XX
KW Human; secreted/transmembrane protein; PRO; tumour; cancer; cytostatic.
XX
OS Homo sapiens.
XX
PN US2003009013-A1.
XX
PD 09-JAN-2003.
XX
PF 01-MAY-2002; 2002US-00063519.
XX
PR 30-DEC-1998; 98KR-00062142.
PR 08-MAR-1999; 99WO-US005028.
PR 14-MAY-1999; 99US-00311832.
PR 14-MAY-1999; 99WO-US010733.
PR 25-AUG-1999; 99US-00380137.
PR 25-AUG-1999; 99US-00380138.
PR 25-AUG-1999; 99US-00380139.
PR 25-AUG-1999; 99US-00380142.
PR 15-SEP-1999; 99US-00397342.
PR 18-OCT-1999; 99US-00403297.
PR 12-NOV-1999; 99US-00423844.
PR 30-DEC-1999; 99WO-US031274.
PR 18-FEB-2000; 2000WO-US004341.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005841.
PR 21-MAR-2000; 2000WO-US007532.
PR 22-MAY-2000; 2000WO-US014042.
PR 02-JUN-2000; 2000WO-US015264.
PR 22-AUG-2000; 2000US-00644848.
PR 24-AUG-2000; 2000WO-US023328.
PR 18-SEP-2000; 2000US-00664610.
PR 18-SEP-2000; 2000US-00665350.
PR 08-NOV-2000; 2000US-00709238.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001WO-US006520.
PR 22-MAR-2001; 2001US-00816744.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 30-MAY-2001; 2001US-00870574.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 29-JUN-2001; 2001US-00869599.
PR 18-JUL-2001; 2001US-00908827.
PR 06-DEC-2001; 2001US-00006867.
XX
PA (GETH) GENENTECH INC.
XX
PI Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Watanabe CK, Wood WI;
XX
DR WPI; 2003-447384/42.
DR N-PSDB; ACD81543.
XX
PT New isolated antibody specifically binding a PRO polypeptide, useful for
PT the preparation of a medicament for treating disorders with the aberrant
PT expression or activity of the PRO polypeptide, such as tumor conditions
PT and cancer.
XX
PS Disclosure; Fig 6; 223pp; English.
XX
CC The invention relates to an antibody that binds to a secreted or
CC transmembrane protein designated PRO1446 appearing as ABO33941. The
CC protein is one of 84 PRO polypeptides which (along with their encoding
CC nucleic acids) are disclosed in the specification. The methods and
CC compositions of the present invention are useful for the preparation of a
CC medicament for the treatment of disorders associated with the aberrant
CC expression or activity of the PRO polypeptide, such as tumour conditions
CC and cancer. They can also be used to generate transgenic or knockout
CC animals useful in the development and screening of therapeutically useful
CC reagents. The PRO polypeptides and encoding nucleic acids can be used as
CC molecular weight markers for protein electrophoresis, chromosome
CC identification and tissue typing. The antibodies may be used in various
CC diagnostic, competitive binding and/or immunoprecipitation assays. The
CC present sequence represents a PRO polypeptide
XX
SQ Sequence 322 AA;

Query Match 100.0%; Score 1657; DB 6; Length 322;
Best Local Similarity 100.0%; Pred. No. 6.1e-148;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
Db 1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

```
Qy          61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy         121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy         181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy         241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy         301 DKNPEESKSPSKTTVRCLEAEV 322
               |||||||||||||||||||||
Db         301 DKNPEESKSPSKTTVRCLEAEV 322
```

Search completed: August 11, 2004, 12:01:05
Job time : 67 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 11, 2004, 11:55:13 ; Search time 47 Seconds
(without alignments)
2150.735 Million cell updates/sec

Title: US-10-063-510-6
Perfect score: 1657
Sequence: 1 MARCFSLVLLLTSIWTTRLL.....NPEESKSPSKTTVRCLEAEV 322

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1292805 seqs, 313927144 residues

Total number of hits satisfying chosen parameters: 1292805

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 500 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1657	100.0	322	9	US-09-909-320-201	Sequence 201, App
2	1657	100.0	322	9	US-09-909-088B-201	Sequence 201, App
3	1657	100.0	322	9	US-09-905-291A-201	Sequence 201, App
4	1657	100.0	322	9	US-09-902-853-201	Sequence 201, App
5	1657	100.0	322	9	US-09-907-824-201	Sequence 201, App
6	1657	100.0	322	9	US-09-907-841-201	Sequence 201, App
7	1657	100.0	322	10	US-09-904-011-201	Sequence 201, App
8	1657	100.0	322	10	US-09-906-742-201	Sequence 201, App
9	1657	100.0	322	10	US-09-906-838-201	Sequence 201, App
10	1657	100.0	322	10	US-09-907-613-201	Sequence 201, App
11	1657	100.0	322	10	US-09-907-942-201	Sequence 201, App
12	1657	100.0	322	10	US-09-904-859-201	Sequence 201, App
13	1657	100.0	322	10	US-09-909-204-201	Sequence 201, App
14	1657	100.0	322	10	US-09-904-820-201	Sequence 201, App
15	1657	100.0	322	10	US-09-904-786-201	Sequence 201, App
16	1657	100.0	322	10	US-09-906-646-201	Sequence 201, App
17	1657	100.0	322	10	US-09-906-700-201	Sequence 201, App
18	1657	100.0	322	10	US-09-903-786-201	Sequence 201, App
19	1657	100.0	322	10	US-09-902-903-201	Sequence 201, App
20	1657	100.0	322	10	US-09-903-749A-201	Sequence 201, App
21	1657	100.0	322	10	US-09-904-119-201	Sequence 201, App
22	1657	100.0	322	10	US-09-904-956-201	Sequence 201, App
23	1657	100.0	322	10	US-09-902-736-201	Sequence 201, App
24	1657	100.0	322	10	US-09-907-794-201	Sequence 201, App
25	1657	100.0	322	10	US-09-903-943-201	Sequence 201, App
26	1657	100.0	322	10	US-09-904-462-201	Sequence 201, App
27	1657	100.0	322	10	US-09-907-925-201	Sequence 201, App
28	1657	100.0	322	10	US-09-902-692-201	Sequence 201, App
29	1657	100.0	322	10	US-09-903-520-201	Sequence 201, App
30	1657	100.0	322	10	US-09-905-056-201	Sequence 201, App
31	1657	100.0	322	10	US-09-909-064-201	Sequence 201, App
32	1657	100.0	322	10	US-09-904-553-201	Sequence 201, App
33	1657	100.0	322	10	US-09-905-381-201	Sequence 201, App
34	1657	100.0	322	10	US-09-905-088-201	Sequence 201, App
35	1657	100.0	322	10	US-09-907-575-201	Sequence 201, App
36	1657	100.0	322	10	US-09-905-075-201	Sequence 201, App
37	1657	100.0	322	10	US-09-902-759-201	Sequence 201, App
38	1657	100.0	322	10	US-09-902-634-201	Sequence 201, App
39	1657	100.0	322	10	US-09-902-713-201	Sequence 201, App
40	1657	100.0	322	10	US-09-907-979-201	Sequence 201, App
41	1657	100.0	322	10	US-09-902-615-201	Sequence 201, App
42	1657	100.0	322	10	US-09-903-925-201	Sequence 201, App
43	1657	100.0	322	10	US-09-906-760A-201	Sequence 201, App
44	1657	100.0	322	10	US-09-903-823-201	Sequence 201, App
45	1657	100.0	322	10	US-09-907-652-201	Sequence 201, App
46	1657	100.0	322	10	US-09-902-572A-201	Sequence 201, App
47	1657	100.0	322	10	US-09-902-979-201	Sequence 201, App
48	1657	100.0	322	10	US-09-905-125-201	Sequence 201, App
49	1657	100.0	322	10	US-09-906-815A-201	Sequence 201, App
50	1657	100.0	322	10	US-09-905-449-201	Sequence 201, App
51	1657	100.0	322	10	US-09-903-806-201	Sequence 201, App
52	1657	100.0	322	10	US-09-904-992-201	Sequence 201, App
53	1657	100.0	322	10	US-09-904-838-201	Sequence 201, App
54	1657	100.0	322	10	US-09-906-777-201	Sequence 201, App
55	1657	100.0	322	10	US-09-903-603A-201	Sequence 201, App
56	1657	100.0	322	10	US-09-904-532-201	Sequence 201, App
57	1657	100.0	322	10	US-09-904-766-201	Sequence 201, App
58	1657	100.0	322	10	US-09-904-920A-201	Sequence 201, App
59	1657	100.0	322	10	US-09-904-877A-201	Sequence 201, App
60	1657	100.0	322	10	US-09-903-562-201	Sequence 201, App
61	1657	100.0	322	10	US-09-906-618-201	Sequence 201, App
62	1657	100.0	322	10	US-09-907-728-201	Sequence 201, App
63	1657	100.0	322	11	US-09-904-805-201	Sequence 201, App
64	1657	100.0	322	11	US-09-904-938A-201	Sequence 201, App
65	1657	100.0	322	11	US-09-906-722A-201	Sequence 201, App
66	1657	100.0	322	11	US-09-908-576-201	Sequence 201, App
67	1657	100.0	322	12	US-10-081-056-32	Sequence 32, Appl
68	1657	100.0	322	12	US-10-063-745-6	Sequence 6, Appli
69	1657	100.0	322	12	US-09-903-640-201	Sequence 201, App
70	1657	100.0	322	12	US-09-908-093-201	Sequence 201, App
71	1657	100.0	322	12	US-10-063-512-6	Sequence 6, Appli
72	1657	100.0	322	12	US-10-063-513-6	Sequence 6, Appli
73	1657	100.0	322	12	US-10-063-515-6	Sequence 6, Appli
74	1657	100.0	322	12	US-10-063-549-6	Sequence 6, Appli
75	1657	100.0	322	12	US-10-063-569-6	Sequence 6, Appli
76	1657	100.0	322	12	US-10-063-551-6	Sequence 6, Appli
77	1657	100.0	322	12	US-09-904-485-201	Sequence 201, App
78	1657	100.0	322	12	US-09-905-348-201	Sequence 201, App
79	1657	100.0	322	12	US-10-063-555-6	Sequence 6, Appli
80	1657	100.0	322	12	US-10-063-563-6	Sequence 6, Appli
81	1657	100.0	322	12	US-10-063-594-6	Sequence 6, Appli
82	1657	100.0	322	12	US-10-063-553-6	Sequence 6, Appli
83	1657	100.0	322	12	US-10-063-554-6	Sequence 6, Appli
84	1657	100.0	322	12	US-10-298-993-201	Sequence 201, App
85	1657	100.0	322	12	US-10-305-654-32	Sequence 32, Appl
86	1657	100.0	322	12	US-10-448-923-201	Sequence 201, App
87	1657	100.0	322	13	US-10-006-867-6	Sequence 6, Appli
88	1657	100.0	322	13	US-10-063-547-6	Sequence 6, Appli

89	1657	100.0	322	14	US-10-063-616-6	Sequence 6, Appli
90	1657	100.0	322	14	US-10-063-502-6	Sequence 6, Appli
91	1657	100.0	322	14	US-10-063-518-6	Sequence 6, Appli
92	1657	100.0	322	14	US-10-063-598-6	Sequence 6, Appli
93	1657	100.0	322	14	US-10-227-693-6	Sequence 6, Appli
94	1657	100.0	322	14	US-10-063-567-6	Sequence 6, Appli
95	1657	100.0	322	14	US-10-063-538-6	Sequence 6, Appli
96	1657	100.0	322	14	US-10-063-599-6	Sequence 6, Appli
97	1657	100.0	322	14	US-10-063-595-6	Sequence 6, Appli
98	1657	100.0	322	14	US-10-223-085-32	Sequence 32, Appl
99	1657	100.0	322	14	US-10-223-084-32	Sequence 32, Appl
100	1657	100.0	322	14	US-10-223-088-32	Sequence 32, Appl
101	1657	100.0	322	14	US-10-223-090-32	Sequence 32, Appl
102	1657	100.0	322	14	US-10-063-580-6	Sequence 6, Appli
103	1657	100.0	322	14	US-10-223-087-32	Sequence 32, Appl
104	1657	100.0	322	14	US-10-063-557-6	Sequence 6, Appli
105	1657	100.0	322	14	US-10-063-585-6	Sequence 6, Appli
106	1657	100.0	322	14	US-10-223-083-32	Sequence 32, Appl
107	1657	100.0	322	14	US-10-079-111-1	Sequence 1, Appli
108	1657	100.0	322	14	US-10-223-089-32	Sequence 32, Appl
109	1657	100.0	322	14	US-10-063-588-6	Sequence 6, Appli
110	1657	100.0	322	14	US-10-063-735-6	Sequence 6, Appli
111	1657	100.0	322	14	US-10-063-526-6	Sequence 6, Appli
112	1657	100.0	322	14	US-10-063-586-6	Sequence 6, Appli
113	1657	100.0	322	14	US-10-299-976-201	Sequence 201, App
114	1657	100.0	322	14	US-10-063-546-6	Sequence 6, Appli
115	1657	100.0	322	14	US-10-063-564-6	Sequence 6, Appli
116	1657	100.0	322	14	US-10-063-662-6	Sequence 6, Appli
117	1657	100.0	322	14	US-10-063-510-6	Sequence 6, Appli
118	1657	100.0	322	14	US-10-063-669-6	Sequence 6, Appli
119	1657	100.0	322	14	US-10-063-670-6	Sequence 6, Appli
120	1657	100.0	322	14	US-10-063-671-6	Sequence 6, Appli
121	1657	100.0	322	14	US-10-063-674-6	Sequence 6, Appli
122	1657	100.0	322	14	US-10-063-675-6	Sequence 6, Appli
123	1657	100.0	322	14	US-10-063-676-6	Sequence 6, Appli
124	1657	100.0	322	14	US-10-063-686-6	Sequence 6, Appli
125	1657	100.0	322	14	US-10-063-689-6	Sequence 6, Appli
126	1657	100.0	322	14	US-10-063-692-6	Sequence 6, Appli
127	1657	100.0	322	14	US-10-063-693-6	Sequence 6, Appli
128	1657	100.0	322	14	US-10-063-694-6	Sequence 6, Appli
129	1657	100.0	322	14	US-10-063-698-6	Sequence 6, Appli
130	1657	100.0	322	14	US-10-063-699-6	Sequence 6, Appli
131	1657	100.0	322	14	US-10-063-702-6	Sequence 6, Appli
132	1657	100.0	322	14	US-10-063-705-6	Sequence 6, Appli
133	1657	100.0	322	14	US-10-063-707-6	Sequence 6, Appli
134	1657	100.0	322	14	US-10-063-709-6	Sequence 6, Appli
135	1657	100.0	322	14	US-10-063-713-6	Sequence 6, Appli
136	1657	100.0	322	14	US-10-063-724-6	Sequence 6, Appli
137	1657	100.0	322	14	US-10-063-727-6	Sequence 6, Appli
138	1657	100.0	322	14	US-10-063-730-6	Sequence 6, Appli
139	1657	100.0	322	14	US-10-063-734-6	Sequence 6, Appli
140	1657	100.0	322	14	US-10-063-736-6	Sequence 6, Appli
141	1657	100.0	322	14	US-10-063-742-6	Sequence 6, Appli
142	1657	100.0	322	14	US-10-063-743-6	Sequence 6, Appli
143	1657	100.0	322	14	US-10-063-744-6	Sequence 6, Appli
144	1657	100.0	322	14	US-10-063-565-6	Sequence 6, Appli
145	1657	100.0	322	14	US-10-063-577-6	Sequence 6, Appli
146	1657	100.0	322	14	US-10-063-591-6	Sequence 6, Appli
147	1657	100.0	322	14	US-10-063-610-6	Sequence 6, Appli
148	1657	100.0	322	14	US-10-063-673-6	Sequence 6, Appli
149	1657	100.0	322	14	US-10-063-710-6	Sequence 6, Appli
150	1657	100.0	322	14	US-10-063-711-6	Sequence 6, Appli
151	1657	100.0	322	14	US-10-063-712-6	Sequence 6, Appli
152	1657	100.0	322	14	US-10-063-714-6	Sequence 6, Appli
153	1657	100.0	322	14	US-10-063-715-6	Sequence 6, Appli
154	1657	100.0	322	14	US-10-063-716-6	Sequence 6, Appli
155	1657	100.0	322	14	US-10-063-717-6	Sequence 6, Appli
156	1657	100.0	322	14	US-10-063-720-6	Sequence 6, Appli
157	1657	100.0	322	14	US-10-063-722-6	Sequence 6, Appli
158	1657	100.0	322	14	US-10-063-726-6	Sequence 6, Appli
159	1657	100.0	322	14	US-10-063-728-6	Sequence 6, Appli
160	1657	100.0	322	14	US-10-063-731-6	Sequence 6, Appli
161	1657	100.0	322	14	US-10-063-732-6	Sequence 6, Appli
162	1657	100.0	322	14	US-10-063-523-6	Sequence 6, Appli
163	1657	100.0	322	14	US-10-063-527-6	Sequence 6, Appli
164	1657	100.0	322	14	US-10-063-579-6	Sequence 6, Appli
165	1657	100.0	322	14	US-10-063-581-6	Sequence 6, Appli
166	1657	100.0	322	14	US-10-063-583-6	Sequence 6, Appli
167	1657	100.0	322	14	US-10-063-589-6	Sequence 6, Appli
168	1657	100.0	322	14	US-10-063-593-6	Sequence 6, Appli
169	1657	100.0	322	14	US-10-063-596-6	Sequence 6, Appli
170	1657	100.0	322	14	US-10-063-600-6	Sequence 6, Appli
171	1657	100.0	322	14	US-10-063-604-6	Sequence 6, Appli
172	1657	100.0	322	14	US-10-063-607-6	Sequence 6, Appli
173	1657	100.0	322	14	US-10-063-612-6	Sequence 6, Appli
174	1657	100.0	322	14	US-10-063-615-6	Sequence 6, Appli
175	1657	100.0	322	14	US-10-063-640-6	Sequence 6, Appli
176	1657	100.0	322	14	US-10-063-642-6	Sequence 6, Appli
177	1657	100.0	322	14	US-10-063-644-6	Sequence 6, Appli
178	1657	100.0	322	14	US-10-063-649-6	Sequence 6, Appli
179	1657	100.0	322	14	US-10-063-650-6	Sequence 6, Appli
180	1657	100.0	322	14	US-10-063-652-6	Sequence 6, Appli
181	1657	100.0	322	14	US-10-063-654-6	Sequence 6, Appli
182	1657	100.0	322	14	US-10-063-659-6	Sequence 6, Appli
183	1657	100.0	322	14	US-10-063-661-6	Sequence 6, Appli
184	1657	100.0	322	14	US-10-063-528-6	Sequence 6, Appli
185	1657	100.0	322	14	US-10-063-540-6	Sequence 6, Appli
186	1657	100.0	322	14	US-10-063-568-6	Sequence 6, Appli
187	1657	100.0	322	14	US-10-063-570-6	Sequence 6, Appli
188	1657	100.0	322	14	US-10-063-582-6	Sequence 6, Appli
189	1657	100.0	322	14	US-10-063-587-6	Sequence 6, Appli
190	1657	100.0	322	14	US-10-063-592-6	Sequence 6, Appli
191	1657	100.0	322	14	US-10-063-597-6	Sequence 6, Appli
192	1657	100.0	322	14	US-10-063-602-6	Sequence 6, Appli
193	1657	100.0	322	14	US-10-063-606-6	Sequence 6, Appli
194	1657	100.0	322	14	US-10-063-609-6	Sequence 6, Appli
195	1657	100.0	322	14	US-10-063-611-6	Sequence 6, Appli
196	1657	100.0	322	14	US-10-063-614-6	Sequence 6, Appli
197	1657	100.0	322	14	US-10-063-639-6	Sequence 6, Appli
198	1657	100.0	322	14	US-10-063-643-6	Sequence 6, Appli
199	1657	100.0	322	14	US-10-063-646-6	Sequence 6, Appli
200	1657	100.0	322	14	US-10-063-651-6	Sequence 6, Appli
201	1657	100.0	322	14	US-10-063-653-6	Sequence 6, Appli
202	1657	100.0	322	14	US-10-063-660-6	Sequence 6, Appli
203	1657	100.0	322	14	US-10-063-665-6	Sequence 6, Appli
204	1657	100.0	322	14	US-10-063-584-6	Sequence 6, Appli
205	1657	100.0	322	14	US-10-063-536-6	Sequence 6, Appli
206	1657	100.0	322	14	US-10-063-562-6	Sequence 6, Appli
207	1657	100.0	322	14	US-10-063-638-6	Sequence 6, Appli
208	1657	100.0	322	14	US-10-063-666-6	Sequence 6, Appli
209	1657	100.0	322	14	US-10-063-672-6	Sequence 6, Appli
210	1657	100.0	322	14	US-10-063-682-6	Sequence 6, Appli
211	1657	100.0	322	14	US-10-063-721-6	Sequence 6, Appli
212	1657	100.0	322	14	US-10-063-723-6	Sequence 6, Appli
213	1657	100.0	322	14	US-10-063-514-6	Sequence 6, Appli
214	1657	100.0	322	14	US-10-063-516-6	Sequence 6, Appli
215	1657	100.0	322	14	US-10-063-529-6	Sequence 6, Appli
216	1657	100.0	322	14	US-10-299-937-201	Sequence 201, App
217	1657	100.0	322	14	US-10-063-684-6	Sequence 6, Appli
218	1657	100.0	322	14	US-10-063-688-6	Sequence 6, Appli
219	1657	100.0	322	14	US-10-223-081-32	Sequence 32, Appl
220	1657	100.0	322	14	US-10-063-520-6	Sequence 6, Appli
221	1657	100.0	322	14	US-10-063-647-6	Sequence 6, Appli
222	1657	100.0	322	14	US-10-063-548-6	Sequence 6, Appli
223	1657	100.0	322	14	US-10-063-578-6	Sequence 6, Appli
224	1657	100.0	322	14	US-10-063-648-6	Sequence 6, Appli
225	1657	100.0	322	14	US-10-063-677-6	Sequence 6, Appli
226	1657	100.0	322	14	US-10-063-718-6	Sequence 6, Appli
227	1657	100.0	322	14	US-10-063-741-6	Sequence 6, Appli
228	1657	100.0	322	14	US-10-063-617-6	Sequence 6, Appli
229	1657	100.0	322	14	US-10-223-082-32	Sequence 32, Appl
230	1657	100.0	322	14	US-10-063-664-6	Sequence 6, Appli
231	1657	100.0	322	14	US-10-063-561-6	Sequence 6, Appli
232	1657	100.0	322	14	US-10-063-618-6	Sequence 6, Appli
233	1657	100.0	322	14	US-10-063-657-6	Sequence 6, Appli
234	1657	100.0	322	14	US-10-063-668-6	Sequence 6, Appli

235	1657	100.0	322	15	US-10-063-550-6	Sequence 6, Appli
236	1657	100.0	322	15	US-10-449-656-201	Sequence 201, App
237	1657	100.0	322	15	US-10-448-713-201	Sequence 201, App
238	1657	100.0	322	16	US-10-425-447-201	Sequence 201, App
239	1657	100.0	322	16	US-10-215-371-201	Sequence 201, App
240	1654	99.8	344	12	US-10-276-774-2349	Sequence 2349, Ap
241	1651	99.6	322	14	US-10-291-634-2	Sequence 2, Appli
242	1651	99.6	322	15	US-10-264-237-2665	Sequence 2665, Ap
243	1511.5	91.2	297	15	US-10-138-588-2	Sequence 2, Appli
244	821	49.5	255	10	US-09-866-050A-700	Sequence 700, App
245	224.5	13.5	339	14	US-10-291-634-3	Sequence 3, Appli
246	222.5	13.4	365	9	US-09-870-759-73	Sequence 73, Appl
247	222.5	13.4	365	10	US-09-751-708A-73	Sequence 73, Appl
248	207	12.5	361	12	US-09-836-544-32	Sequence 32, Appl
249	203	12.3	742	16	US-10-648-593-176	Sequence 176, App
250	202	12.2	361	16	US-10-663-244-3	Sequence 3, Appli
251	202	12.2	742	10	US-09-983-000A-16	Sequence 16, Appl
252	202	12.2	742	15	US-10-116-275-206	Sequence 206, App
253	202	12.2	742	16	US-10-663-244-1	Sequence 1, Appli
254	199.5	12.0	493	12	US-09-836-544-34	Sequence 34, Appl
255	199.5	12.0	493	16	US-10-663-244-2	Sequence 2, Appli
256	199.5	12.0	700	14	US-10-012-969C-2	Sequence 2, Appli
257	170	10.3	34	9	US-09-864-761-47905	Sequence 47905, A
258	161	9.7	90	9	US-09-799-118-3	Sequence 3, Appli
259	153.5	9.3	90	10	US-09-927-463-8	Sequence 8, Appli
260	150	9.1	510	15	US-10-104-047-2580	Sequence 2580, Ap
261	145	8.8	277	9	US-09-799-118-2	Sequence 2, Appli
262	145	8.8	277	14	US-10-247-671-185	Sequence 185, App
263	145	8.8	277	15	US-10-295-027-292	Sequence 292, App
264	145	8.8	277	15	US-10-295-027-1237	Sequence 1237, Ap
265	145	8.8	1069	15	US-10-028-248A-39	Sequence 39, Appl
266	145	8.8	1069	15	US-10-107-782-39	Sequence 39, Appl
267	145	8.8	2626	12	US-10-634-574-4	Sequence 4, Appli
268	143.5	8.7	95	12	US-10-424-599-248857	Sequence 248857,
269	134	8.1	897	15	US-10-028-248A-41	Sequence 41, Appl
270	134	8.1	897	15	US-10-028-248A-211	Sequence 211, App
271	134	8.1	897	15	US-10-107-782-41	Sequence 41, Appl
272	134	8.1	897	15	US-10-107-782-211	Sequence 211, App
273	134	8.1	2675	15	US-10-028-248A-2	Sequence 2, Appli
274	134	8.1	2675	15	US-10-107-782-2	Sequence 2, Appli
275	133	8.0	1192	15	US-10-028-248A-40	Sequence 40, Appl
276	133	8.0	1192	15	US-10-107-782-40	Sequence 40, Appl
277	133	8.0	1394	9	US-09-842-930A-25	Sequence 25, Appl
278	133	8.0	1416	14	US-10-133-172-4	Sequence 4, Appli
279	133	8.0	1431	9	US-09-842-930A-2	Sequence 2, Appli
280	133	8.0	1431	14	US-10-133-172-2	Sequence 2, Appli
281	133	8.0	1653	14	US-10-133-172-20	Sequence 20, Appl
282	131.5	7.9	883	12	US-10-453-420-3	Sequence 3, Appli
283	131.5	7.9	883	14	US-10-195-970-3	Sequence 3, Appli
284	131	7.9	649	10	US-09-759-130B-333	Sequence 333, App
285	131	7.9	649	14	US-10-189-123-63	Sequence 63, Appl
286	131	7.9	649	14	US-10-188-495-63	Sequence 63, Appl
287	131	7.9	649	16	US-10-741-790-333	Sequence 333, App
288	131	7.9	671	10	US-09-759-130B-331	Sequence 331, App
289	131	7.9	671	14	US-10-189-123-61	Sequence 61, Appl
290	131	7.9	671	14	US-10-188-495-61	Sequence 61, Appl
291	131	7.9	671	16	US-10-741-790-331	Sequence 331, App
292	131	7.9	911	12	US-10-312-352-24	Sequence 24, Appl
293	131	7.9	911	12	US-10-245-752-52	Sequence 52, Appl
294	131	7.9	911	12	US-10-245-859-52	Sequence 52, Appl
295	131	7.9	911	14	US-10-245-103-52	Sequence 52, Appl
296	131	7.9	911	14	US-10-245-107-52	Sequence 52, Appl
297	131	7.9	911	14	US-10-245-143-52	Sequence 52, Appl
298	131	7.9	911	14	US-10-245-771-52	Sequence 52, Appl
299	131	7.9	911	14	US-10-245-851-52	Sequence 52, Appl
300	131	7.9	911	14	US-10-245-883-52	Sequence 52, Appl
301	131	7.9	911	14	US-10-237-535-52	Sequence 52, Appl
302	131	7.9	911	14	US-10-238-183-52	Sequence 52, Appl
303	131	7.9	911	14	US-10-238-283-52	Sequence 52, Appl
304	131	7.9	911	14	US-10-238-370-52	Sequence 52, Appl
305	131	7.9	911	14	US-10-245-055-52	Sequence 52, Appl
306	131	7.9	911	14	US-10-245-147-52	Sequence 52, Appl
307	131	7.9	911	14	US-10-245-730-52	Sequence 52, Appl
308	131	7.9	911	14	US-10-245-739-52	Sequence 52, Appl
309	131	7.9	911	14	US-10-246-210-52	Sequence 52, Appl
310	131	7.9	911	14	US-10-239-196-52	Sequence 52, Appl
311	131	7.9	911	14	US-10-243-024-52	Sequence 52, Appl
312	131	7.9	911	14	US-10-243-409-52	Sequence 52, Appl
313	131	7.9	911	14	US-10-245-621-52	Sequence 52, Appl
314	131	7.9	911	14	US-10-245-880-52	Sequence 52, Appl
315	131	7.9	911	14	US-10-245-033-52	Sequence 52, Appl
316	131	7.9	911	14	US-10-243-095-52	Sequence 52, Appl
317	131	7.9	911	14	US-10-245-185-52	Sequence 52, Appl
318	131	7.9	911	14	US-10-245-427-52	Sequence 52, Appl
319	131	7.9	911	14	US-10-245-473-52	Sequence 52, Appl
320	131	7.9	911	14	US-10-245-770-52	Sequence 52, Appl
321	131	7.9	911	14	US-10-245-877-52	Sequence 52, Appl
322	131	7.9	911	14	US-10-246-976-52	Sequence 52, Appl
323	131	7.9	911	14	US-10-243-320-52	Sequence 52, Appl
324	131	7.9	911	14	US-10-242-743-52	Sequence 52, Appl
325	131	7.9	911	14	US-10-242-845-52	Sequence 52, Appl
326	131	7.9	911	14	US-10-237-636-52	Sequence 52, Appl
327	131	7.9	911	14	US-10-238-325-52	Sequence 52, Appl
328	131	7.9	911	14	US-10-238-346-52	Sequence 52, Appl
329	131	7.9	911	14	US-10-238-411-52	Sequence 52, Appl
330	131	7.9	911	14	US-10-243-124-52	Sequence 52, Appl
331	131	7.9	911	14	US-10-243-425-52	Sequence 52, Appl
332	131	7.9	911	14	US-10-243-446-52	Sequence 52, Appl
333	131	7.9	911	14	US-10-245-874-52	Sequence 52, Appl
334	131	7.9	911	14	US-10-242-653-52	Sequence 52, Appl
335	131	7.9	911	14	US-10-243-167-52	Sequence 52, Appl
336	131	7.9	911	14	US-10-243-388-52	Sequence 52, Appl
337	131	7.9	911	14	US-10-244-947-52	Sequence 52, Appl
338	131	7.9	911	14	US-10-244-968-52	Sequence 52, Appl
339	131	7.9	911	14	US-10-244-990-52	Sequence 52, Appl
340	131	7.9	911	14	US-10-245-079-52	Sequence 52, Appl
341	131	7.9	911	14	US-10-245-127-52	Sequence 52, Appl
342	131	7.9	911	14	US-10-245-207-52	Sequence 52, Appl
343	131	7.9	911	14	US-10-245-646-52	Sequence 52, Appl
344	131	7.9	911	14	US-10-245-695-52	Sequence 52, Appl
345	131	7.9	911	14	US-10-245-699-52	Sequence 52, Appl
346	131	7.9	911	14	US-10-245-737-52	Sequence 52, Appl
347	131	7.9	911	14	US-10-245-878-52	Sequence 52, Appl
348	131	7.9	911	14	US-10-245-890-52	Sequence 52, Appl
349	131	7.9	911	14	US-10-245-899-52	Sequence 52, Appl
350	131	7.9	911	14	US-10-245-900-52	Sequence 52, Appl
351	131	7.9	911	14	US-10-247-058-52	Sequence 52, Appl
352	131	7.9	911	14	US-10-245-454-52	Sequence 52, Appl
353	131	7.9	911	14	US-10-237-471-52	Sequence 52, Appl
354	131	7.9	911	14	US-10-238-261-52	Sequence 52, Appl
355	131	7.9	911	14	US-10-238-324-52	Sequence 52, Appl
356	131	7.9	911	14	US-10-241-860-52	Sequence 52, Appl
357	131	7.9	911	14	US-10-242-172-52	Sequence 52, Appl
358	131	7.9	911	14	US-10-242-652-52	Sequence 52, Appl
359	131	7.9	911	14	US-10-242-990-52	Sequence 52, Appl
360	131	7.9	911	14	US-10-243-023-52	Sequence 52, Appl
361	131	7.9	911	14	US-10-243-103-52	Sequence 52, Appl
362	131	7.9	911	14	US-10-243-276-52	Sequence 52, Appl
363	131	7.9	911	14	US-10-243-326-52	Sequence 52, Appl
364	131	7.9	911	14	US-10-243-364-52	Sequence 52, Appl
365	131	7.9	911	14	US-10-243-494-52	Sequence 52, Appl
366	131	7.9	911	14	US-10-244-995-52	Sequence 52, Appl
367	131	7.9	911	14	US-10-245-230-52	Sequence 52, Appl
368	131	7.9	911	14	US-10-245-253-52	Sequence 52, Appl
369	131	7.9	911	14	US-10-245-479-52	Sequence 52, Appl
370	131	7.9	911	14	US-10-245-499-52	Sequence 52, Appl
371	131	7.9	911	14	US-10-245-772-52	Sequence 52, Appl
372	131	7.9	911	14	US-10-245-811-52	Sequence 52, Appl
373	131	7.9	911	14	US-10-245-812-52	Sequence 52, Appl
374	131	7.9	911	14	US-10-245-852-52	Sequence 52, Appl
375	131	7.9	911	14	US-10-245-875-52	Sequence 52, Appl
376	131	7.9	911	14	US-10-245-881-52	Sequence 52, Appl
377	131	7.9	911	14	US-10-245-911-52	Sequence 52, Appl
378	131	7.9	911	14	US-10-245-913-52	Sequence 52, Appl
379	131	7.9	911	14	US-10-246-080-52	Sequence 52, Appl
380	131	7.9	911	14	US-10-246-121-52	Sequence 52, Appl

381	131	7.9	911	14	US-10-246-305-52	Sequence 52, Appl
382	131	7.9	911	14	US-10-246-929-52	Sequence 52, Appl
383	131	7.9	911	14	US-10-247-036-52	Sequence 52, Appl
384	131	7.9	911	14	US-10-243-255-52	Sequence 52, Appl
385	131	7.9	911	14	US-10-245-810-52	Sequence 52, Appl
386	131	7.9	911	14	US-10-245-910-52	Sequence 52, Appl
387	131	7.9	911	14	US-10-246-098-52	Sequence 52, Appl
388	131	7.9	911	14	US-10-237-496-52	Sequence 52, Appl
389	131	7.9	911	14	US-10-242-074-52	Sequence 52, Appl
390	131	7.9	911	14	US-10-242-505-52	Sequence 52, Appl
391	131	7.9	911	14	US-10-242-574-52	Sequence 52, Appl
392	131	7.9	911	14	US-10-243-261-52	Sequence 52, Appl
393	131	7.9	911	14	US-10-243-282-52	Sequence 52, Appl
394	131	7.9	911	14	US-10-243-402-52	Sequence 52, Appl
395	131	7.9	911	14	US-10-243-431-52	Sequence 52, Appl
396	131	7.9	911	14	US-10-245-164-52	Sequence 52, Appl
397	131	7.9	911	14	US-10-244-972-52	Sequence 52, Appl
398	131	7.9	911	14	US-10-197-942-52	Sequence 52, Appl
399	131	7.9	911	14	US-10-238-196-52	Sequence 52, Appl
400	131	7.9	911	14	US-10-245-013-52	Sequence 52, Appl
401	130.5	7.9	883	10	US-09-759-130B-342	Sequence 342, App
402	130.5	7.9	883	12	US-10-453-420-6	Sequence 6, Appli
403	130.5	7.9	883	14	US-10-195-970-6	Sequence 6, Appli
404	130.5	7.9	883	14	US-10-189-123-72	Sequence 72, Appl
405	130.5	7.9	883	14	US-10-188-495-72	Sequence 72, Appl
406	130.5	7.9	883	16	US-10-741-790-342	Sequence 342, App
407	130.5	7.9	1082	15	US-10-264-237-2725	Sequence 2725, Ap
408	129.5	7.8	2570	15	US-10-028-248A-42	Sequence 42, Appl
409	129.5	7.8	2570	15	US-10-107-782-42	Sequence 42, Appl
410	128	7.7	2212	15	US-10-028-248A-43	Sequence 43, Appl
411	128	7.7	2212	15	US-10-107-782-43	Sequence 43, Appl
412	127.5	7.7	671	10	US-09-983-000A-14	Sequence 14, Appl
413	127.5	7.7	911	12	US-10-453-420-8	Sequence 8, Appli
414	127.5	7.7	911	15	US-10-295-027-260	Sequence 260, App
415	127	7.7	277	15	US-10-295-027-294	Sequence 294, App
416	126.5	7.6	457	10	US-09-774-639-108	Sequence 108, App
417	126.5	7.6	457	10	US-09-969-730-110	Sequence 110, App
418	126.5	7.6	457	16	US-10-621-363-110	Sequence 110, App
419	124.5	7.5	482	15	US-10-104-047-3794	Sequence 3794, Ap
420	121.5	7.3	360	9	US-09-909-320-213	Sequence 213, App
421	121.5	7.3	360	9	US-09-909-088B-213	Sequence 213, App
422	121.5	7.3	360	9	US-09-905-291A-213	Sequence 213, App
423	121.5	7.3	360	9	US-09-902-853-213	Sequence 213, App
424	121.5	7.3	360	9	US-09-907-824-213	Sequence 213, App
425	121.5	7.3	360	9	US-09-907-841-213	Sequence 213, App
426	121.5	7.3	360	10	US-09-904-011-213	Sequence 213, App
427	121.5	7.3	360	10	US-09-906-742-213	Sequence 213, App
428	121.5	7.3	360	10	US-09-906-838-213	Sequence 213, App
429	121.5	7.3	360	10	US-09-907-613-213	Sequence 213, App
430	121.5	7.3	360	10	US-09-907-942-213	Sequence 213, App
431	121.5	7.3	360	10	US-09-904-859-213	Sequence 213, App
432	121.5	7.3	360	10	US-09-909-204-213	Sequence 213, App
433	121.5	7.3	360	10	US-09-904-820-213	Sequence 213, App
434	121.5	7.3	360	10	US-09-904-786-213	Sequence 213, App
435	121.5	7.3	360	10	US-09-906-646-213	Sequence 213, App
436	121.5	7.3	360	10	US-09-906-700-213	Sequence 213, App
437	121.5	7.3	360	10	US-09-903-786-213	Sequence 213, App
438	121.5	7.3	360	10	US-09-902-903-213	Sequence 213, App
439	121.5	7.3	360	10	US-09-903-749A-213	Sequence 213, App
440	121.5	7.3	360	10	US-09-904-119-213	Sequence 213, App
441	121.5	7.3	360	10	US-09-904-956-213	Sequence 213, App
442	121.5	7.3	360	10	US-09-902-736-213	Sequence 213, App
443	121.5	7.3	360	10	US-09-907-794-213	Sequence 213, App
444	121.5	7.3	360	10	US-09-903-943-213	Sequence 213, App
445	121.5	7.3	360	10	US-09-904-462-213	Sequence 213, App
446	121.5	7.3	360	10	US-09-907-925-213	Sequence 213, App
447	121.5	7.3	360	10	US-09-902-692-213	Sequence 213, App
448	121.5	7.3	360	10	US-09-903-520-213	Sequence 213, App
449	121.5	7.3	360	10	US-09-905-056-213	Sequence 213, App
450	121.5	7.3	360	10	US-09-909-064-213	Sequence 213, App
451	121.5	7.3	360	10	US-09-904-553-213	Sequence 213, App
452	121.5	7.3	360	10	US-09-905-381-213	Sequence 213, App
453	121.5	7.3	360	10	US-09-905-088-213	Sequence 213, App
454	121.5	7.3	360	10	US-09-907-575-213	Sequence 213, App
455	121.5	7.3	360	10	US-09-905-075-213	Sequence 213, App
456	121.5	7.3	360	10	US-09-902-759-213	Sequence 213, App
457	121.5	7.3	360	10	US-09-902-634-213	Sequence 213, App
458	121.5	7.3	360	10	US-09-902-713-213	Sequence 213, App
459	121.5	7.3	360	10	US-09-907-979-213	Sequence 213, App
460	121.5	7.3	360	10	US-09-902-615-213	Sequence 213, App
461	121.5	7.3	360	10	US-09-903-925-213	Sequence 213, App
462	121.5	7.3	360	10	US-09-906-760A-213	Sequence 213, App
463	121.5	7.3	360	10	US-09-903-823-213	Sequence 213, App
464	121.5	7.3	360	10	US-09-907-652-213	Sequence 213, App
465	121.5	7.3	360	10	US-09-902-572A-213	Sequence 213, App
466	121.5	7.3	360	10	US-09-902-979-213	Sequence 213, App
467	121.5	7.3	360	10	US-09-905-125-213	Sequence 213, App
468	121.5	7.3	360	10	US-09-906-815A-213	Sequence 213, App
469	121.5	7.3	360	10	US-09-905-449-213	Sequence 213, App
470	121.5	7.3	360	10	US-09-903-806-213	Sequence 213, App
471	121.5	7.3	360	10	US-09-904-992-213	Sequence 213, App
472	121.5	7.3	360	10	US-09-904-838-213	Sequence 213, App
473	121.5	7.3	360	10	US-09-906-777-213	Sequence 213, App
474	121.5	7.3	360	10	US-09-903-603A-213	Sequence 213, App
475	121.5	7.3	360	10	US-09-904-532-213	Sequence 213, App
476	121.5	7.3	360	10	US-09-904-766-213	Sequence 213, App
477	121.5	7.3	360	10	US-09-904-920A-213	Sequence 213, App
478	121.5	7.3	360	10	US-09-904-877A-213	Sequence 213, App
479	121.5	7.3	360	10	US-09-903-562-213	Sequence 213, App
480	121.5	7.3	360	10	US-09-906-618-213	Sequence 213, App
481	121.5	7.3	360	10	US-09-907-728-213	Sequence 213, App
482	121.5	7.3	360	11	US-09-904-805-213	Sequence 213, App
483	121.5	7.3	360	11	US-09-904-938A-213	Sequence 213, App
484	121.5	7.3	360	11	US-09-906-722A-213	Sequence 213, App
485	121.5	7.3	360	11	US-09-908-576-213	Sequence 213, App
486	121.5	7.3	360	12	US-10-206-915-28	Sequence 28, Appl
487	121.5	7.3	360	12	US-10-199-670-28	Sequence 28, Appl
488	121.5	7.3	360	12	US-10-201-858-28	Sequence 28, Appl
489	121.5	7.3	360	12	US-10-205-890-28	Sequence 28, Appl
490	121.5	7.3	360	12	US-10-208-024-28	Sequence 28, Appl
491	121.5	7.3	360	12	US-10-201-853-28	Sequence 28, Appl
492	121.5	7.3	360	12	US-09-903-640-213	Sequence 213, App
493	121.5	7.3	360	12	US-09-908-093-213	Sequence 213, App
494	121.5	7.3	360	12	US-10-174-581-28	Sequence 28, Appl
495	121.5	7.3	360	12	US-10-176-483-28	Sequence 28, Appl
496	121.5	7.3	360	12	US-10-176-749-28	Sequence 28, Appl
497	121.5	7.3	360	12	US-10-176-914-28	Sequence 28, Appl
498	121.5	7.3	360	12	US-10-176-915-28	Sequence 28, Appl
499	121.5	7.3	360	12	US-09-904-485-213	Sequence 213, App
500	121.5	7.3	360	15	US-10-205-506-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1
US-09-909-320-201
; Sequence 201, Application US/09909320
; Patent No. US20020132240A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.

; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/909,320
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 201
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic protein
US-09-909-320-201

Query Match 100.0%; Score 1657; DB 9; Length 322;
Best Local Similarity 100.0%; Pred. No. 1.7e-150;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy        301 DKNPEESKSPSKTTVRCLEAEV 322
              ||||||||||||||||||||||
Db        301 DKNPEESKSPSKTTVRCLEAEV 322
```

RESULT 2
US-09-909-088B-201
; Sequence 201, Application US/09909088B
; Patent No. US20020146709A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/909,088B
; CURRENT FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565

```
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 201
;   LENGTH: 322
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic protein
US-09-909-088B-201

  Query Match             100.0%;  Score 1657;  DB 9;  Length 322;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-150;
  Matches  322;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy        301 DKNPEESKSPSKTTVRCLEAEV 322
              ||||||||||||||||||||||
Db        301 DKNPEESKSPSKTTVRCLEAEV 322


RESULT 3
US-09-905-291A-201
; Sequence 201, Application US/09905291A
; Patent No. US20020160374A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/905,291A
;  CURRENT FILING DATE:  2001-07-12
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 201
;   LENGTH: 322
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic protein
US-09-905-291A-201

  Query Match             100.0%;  Score 1657;  DB 9;  Length 322;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-150;
  Matches  322;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy        301 DKNPEESKSPSKTTVRCLEAEV 322
              ||||||||||||||||||||||
Db        301 DKNPEESKSPSKTTVRCLEAEV 322


RESULT 4
US-09-902-853-201
; Sequence 201, Application US/09902853
; Publication No. US20020192659A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/902,853
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: US/09/665,350
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 201
;   LENGTH: 322
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-09-902-853-201

  Query Match           100.0%;  Score 1657;  DB 9;  Length 322;
  Best Local Similarity  100.0%;  Pred. No. 1.7e-150;
  Matches  322;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEPIVSDSTYSVASPYSTIPA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEPIVSDSTYSVASPYSTIPA 180

Qy        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy        301 DKNPEESKSPSKTTVRCLEAEV 322
              ||||||||||||||||||||||
Db        301 DKNPEESKSPSKTTVRCLEAEV 322


RESULT 5
US-09-907-824-201
; Sequence 201, Application US/09907824
; Publication No. US20020197671A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
```

```
;  CURRENT APPLICATION NUMBER: US/09/907,824
;  CURRENT FILING DATE:  2001-07-17
;  PRIOR APPLICATION NUMBER: 09/665,350
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 201
;   LENGTH: 322
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-09-907-824-201

  Query Match           100.0%;  Score 1657;  DB 9;  Length 322;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-150;
  Matches  322;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy        301 DKNPEESKSPSKTTVRCLEAEV 322
              |||||||||||||||||||||||
Db        301 DKNPEESKSPSKTTVRCLEAEV 322
```

```
RESULT 6
US-09-907-841-201
; Sequence 201, Application US/09907841
; Publication No. US20020198366A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/907,841
;  CURRENT FILING DATE:  2001-11-20
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 201
;   LENGTH: 322
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic protein
US-09-907-841-201

  Query Match           100.0%;  Score 1657;  DB 9;  Length 322;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-150;
  Matches  322;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
```

```
Qy          61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy         121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy         181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPPFVENKAAFKNEAAGFGGVPTAL 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy         241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy         301 DKNPEESKSPSKTTVRCLEAEV 322
               |||||||||||||||||||||
Db         301 DKNPEESKSPSKTTVRCLEAEV 322
```

RESULT 7
US-09-904-011-201
; Sequence 201, Application US/09904011
; Publication No. US20030003530A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/904,011
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 201
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-904-011-201

Query Match 100.0%; Score 1657; DB 10; Length 322;
Best Local Similarity 100.0%; Pred. No. 1.7e-150;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy           1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy          61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy         121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy         181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPPFVENKAAFKNEAAGFGGVPTAL 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy         241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy         301 DKNPEESKSPSKTTVRCLEAEV 322
               |||||||||||||||||||||
Db         301 DKNPEESKSPSKTTVRCLEAEV 322
```

RESULT 8
US-09-906-742-201
; Sequence 201, Application US/09906742
; Publication No. US20030023054A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.

; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/906,742
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 201
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-906-742-201

Query Match 100.0%; Score 1657; DB 10; Length 322;
Best Local Similarity 100.0%; Pred. No. 1.7e-150;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy        301 DKNPEESKSPSKTTVRCLEAEV 322
              ||||||||||||||||||||||
Db        301 DKNPEESKSPSKTTVRCLEAEV 322
```

RESULT 9
US-09-906-838-201
; Sequence 201, Application US/09906838
; Publication No. US20030027143A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/906,838
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214

```
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 201
;   LENGTH: 322
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-09-906-838-201

  Query Match            100.0%;  Score 1657;  DB 10;  Length 322;
  Best Local Similarity  100.0%;  Pred. No. 1.7e-150;
  Matches  322;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy        301 DKNPEESKSPSKTTVRCLEAEV 322
              |||||||||||||||||||||
Db        301 DKNPEESKSPSKTTVRCLEAEV 322


RESULT 10
US-09-907-613-201
; Sequence 201, Application US/09907613
; Publication No. US20030027145A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/907,613
;  CURRENT FILING DATE:  2001-07-17
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 201
;   LENGTH: 322
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic protein
US-09-907-613-201

  Query Match            100.0%;  Score 1657;  DB 10;  Length 322;
  Best Local Similarity  100.0%;  Pred. No. 1.7e-150;
  Matches  322;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy        301 DKNPEESKSPSKTTVRCLEAEV 322
                |||||||||||||||||||||
Db        301 DKNPEESKSPSKTTVRCLEAEV 322


RESULT 11
US-09-907-942-201
; Sequence 201, Application US/09907942
; Publication No. US20030027146A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/907,942
;  CURRENT FILING DATE:  2002-01-22
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 201
;   LENGTH: 322
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic protein
US-09-907-942-201

  Query Match           100.0%;  Score 1657;  DB 10;  Length 322;
  Best Local Similarity  100.0%;  Pred. No. 1.7e-150;
  Matches  322;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy        301 DKNPEESKSPSKTTVRCLEAEV 322
              |||||||||||||||||||||
Db        301 DKNPEESKSPSKTTVRCLEAEV 322


RESULT 12
US-09-904-859-201
; Sequence 201, Application US/09904859
; Publication No. US20030036060A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
```

```
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/904,859
;  CURRENT FILING DATE:  2001-07-12
;  PRIOR APPLICATION NUMBER: 09/665,350
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 201
;   LENGTH: 322
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-09-904-859-201

  Query Match           100.0%;  Score 1657;  DB 10;  Length 322;
  Best Local Similarity  100.0%;  Pred. No. 1.7e-150;
  Matches  322;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy        301 DKNPEESKSPSKTTVRCLEAEV 322
              ||||||||||||||||||||||
Db        301 DKNPEESKSPSKTTVRCLEAEV 322


RESULT 13
US-09-909-204-201
; Sequence 201, Application US/09909204
; Publication No. US20030036061A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/909,204
;  CURRENT FILING DATE:  2001-07-18
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
```

```
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 201
;   LENGTH: 322
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic protein
US-09-909-204-201

  Query Match             100.0%;  Score 1657;  DB 10;  Length 322;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-150;
  Matches  322;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy        301 DKNPEESKSPSKTTVRCLEAEV 322
              |||||||||||||||||||||||
Db        301 DKNPEESKSPSKTTVRCLEAEV 322


RESULT 14
US-09-904-820-201
; Sequence 201, Application US/09904820
; Publication No. US20030036094A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/904,820
;  CURRENT FILING DATE:  2001-07-13
;  PRIOR APPLICATION NUMBER: 09/665,350
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 201
;   LENGTH: 322
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-09-904-820-201

  Query Match             100.0%;  Score 1657;  DB 10;  Length 322;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-150;
  Matches  322;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy        301 DKNPEESKSPSKTTVRCLEAEV 322
              |||||||||||||||||||||||
Db        301 DKNPEESKSPSKTTVRCLEAEV 322
```

```
RESULT 15
US-09-904-786-201
; Sequence 201, Application US/09904786
; Publication No. US20030039969A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/904,786
;  CURRENT FILING DATE:  2001-07-12
;  PRIOR APPLICATION NUMBER: 09/665,350
;  PRIOR FILING DATE: 2000-09-18
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 201
;   LENGTH: 322
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-09-904-786-201

  Query Match            100.0%;  Score 1657;  DB 10;  Length 322;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-150;
  Matches  322;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy        301 DKNPEESKSPSKTTVRCLEAEV 322
              ||||||||||||||||||||||
Db        301 DKNPEESKSPSKTTVRCLEAEV 322
```

Search completed: August 11, 2004, 12:01:54
Job time : 51 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 11, 2004, 11:50:22 ; Search time 17 Seconds
(without alignments)
1821.981 Million cell updates/sec

Title: US-10-063-510-6
Perfect score: 1657
Sequence: 1 MARCFSLVLLLTSIWTTRLL.....NPEESKSPSKTTVRCLEAEV 322

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	229.5	13.9	363	2	A37009	CD44 homolog membr
2	222.5	13.4	362	2	A30901	lymphocyte adhesio
3	222.5	13.4	365	2	A34424	CD44 membrane glyc
4	222	13.4	503	2	B38745	cell adhesion mole
5	217.5	13.1	362	2	A35616	T-cell surface gly
6	208.5	12.6	366	2	A53286	cell-surface glyco
7	205	12.4	361	2	JH0417	cell adhesion mole
8	202.5	12.2	351	2	S45305	CD44 antigen precu
9	202	12.2	742	2	A47195	lymphocyte homing
10	199.5	12.0	426	2	JH0518	lymphocyte homing
11	199.5	12.0	493	2	S13530	CD44E protein, epi
12	198	11.9	395	2	I77371	CD44R5 - human
13	196.5	11.9	699	2	I37369	epican - human
14	186	11.2	359	2	S24240	lymphocyte surface
15	148.5	9.0	2109	1	I50421	aggrecan precursor
16	145.5	8.8	3562	2	A47171	chondroitin sulfat
17	145	8.8	277	2	A41735	hyaluronate-bindin
18	145	8.8	1069	2	T42681	hypothetical prote
19	144	8.7	275	2	JC6506	tumor necrosis fac
20	143	8.6	2327	2	T42630	aggrecan - bovine
21	137	8.3	276	2	A47290	TSG-6 homolog PS4
22	131.5	7.9	2124	2	A28452	proteoglycan core
23	131	7.9	2132	1	A55182	aggrecan precursor
24	130.5	7.9	883	2	S49126	brevican precursor
25	130.5	7.9	883	2	S57653	brevican precursor
26	130.5	7.9	2415	1	A39086	aggrecan precursor
27	128.5	7.8	912	2	A54423	brevican precursor
28	121	7.3	1340	2	A39808	proteoglycan core
29	117	7.1	340	2	JC7505	brain link protein
30	116	7.0	2397	1	A55535	versican precursor
31	115	6.9	370	2	S29139	aggrecan - pig (fr
32	115	6.9	1643	2	T14274	versican precursor
33	115	6.9	3381	2	T42389	versican precursor
34	114.5	6.9	1257	2	S28764	neurocan precursor
35	112	6.8	390	1	QQBE77	glycoprotein I pre
36	111.5	6.7	402	2	S73773	dihydrolipoamide a
37	111.5	6.7	909	1	QRXLL2	LDL receptor 2 pre
38	111	6.7	2409	1	A60979	versican precursor
39	108	6.5	354	1	LKHU	proteoglycan link
40	108	6.5	645	2	T29818	hypothetical prote
41	107.5	6.5	739	2	T21769	hypothetical prote
42	107	6.5	355	1	LKCH	proteoglycan link
43	107	6.5	371	2	A53908	brevican precursor
44	107	6.5	408	1	LKRT2	proteoglycan link
45	107	6.5	1268	2	S52781	neurocan - mouse

ALIGNMENTS

RESULT 1
A37009
CD44 homolog membrane glycoprotein precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 31-Jan-1992 #sequence_revision 31-Jan-1992 #text_change 21-Jul-2000
C;Accession: A37009
R;Zhou, D.F.H.; Ding, J.F.; Picker, L.J.; Bargatze, R.F.; Butcher, E.C.; Goeddel, D.V.
J. Immunol. 143, 3390-3395, 1989
A;Title: Molecular cloning and expression of Pgp-1. The mouse homolog of the human H-CAM
A;Reference number: A37009; MUID:90038499; PMID:2681416
A;Accession: A37009
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-363 <ZHO>
A;Cross-references: GB:M30655; NID:g200332; PIDN:AAA39922.1; PID:g200333
C;Superfamily: human cell adhesion protein CD44
C;Keywords: cell adhesion; glycoprotein; membrane protein

Query Match 13.9%; Score 229.5; DB 2; Length 363;
Best Local Similarity 24.9%; Pred. No. 6.7e-11;
Matches 89; Conservative 57; Mismatches 148; Indels 63; Gaps 14;

```
Qy         15 WTTR---LLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGK 71
              | |      |:| ||  ::: : |:||  |:  | |     :: ||| : |:    :|
Db          6 WHTAWGLCLLQLSLAHQQIDLNVTCRYAGVFHVEKNGRYSISRTEAADLCQAFNSTLPTM 65

Qy         72 DQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNS 131
              ||:: ||   |||| ||:: :| ||| || ||  |  |  || |     :  :  ||:|:
Db         66 DQMKLALSKGFETCRYGFI-EGNVVIPRIHPNAICAANHTGVYILVTSNTSHYDTYCFNA 124

Qy        132 SDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDST-YSVASPYST----IPAPT---- 182
              |      |     |:   : |:     |   :  | | ||    | |    | |
Db        125 SAPPEEDC-----TSVTDLPNSFDGPVTITIVNRDGTRYSKKGEYRTHQEDIDASNIIDD 179

Qy        183 ----------TTPPA-PASTSIPRRKKLICVTEVFMETSTMST-ETEPFVENKAAFK--- 227
                       :||      | :|  :      : |   ||::| : :   ::: : :
Db        180 DVSSGSTIEKSTPEGYILHTYLPTEQPTGDQDDSFFIRSTLATRDRDSSKDSRGSSRTVT 239

Qy        228 --NEAAGFGG-----------------VPTALLVLALLFFGAAAGLGFCYVKRYVKAFP 267
                :| ||                   :|  |::|| |     |  |  |     :
Db        240 HGSELAGHSSANQDSGVTTTSGPMRRPQIPEWLIILASL-LALALILAVC-----IAVNS 293

Qy        268 FTNKNQQKEMI---ETKVVKEEKANDSNPNEESKKTDKNPEESKSPSKTTVRCLEAE 321
                 |:|:::      |:: | :: |  | ||  :    :| ||:|  :|: |:
Db        294 RRRCGQKKKLVINGGNGTVEDRKPSELN-GEASKSQEMVHLVNKEPSETPDQCMTAD 349
```

RESULT 2
A30901
lymphocyte adhesion receptor precursor - baboon
C;Species: Papio sp. (baboon)
C;Date: 18-Apr-1989 #sequence_revision 18-Apr-1989 #text_change 21-Jul-2000

C;Accession: A33935; A30901
R;Idzerda, R.L.; Carter, W.G.; Nottenburg, C.; Wayner, E.A.; Gallatin, W.M.; St. John, T
Proc. Natl. Acad. Sci. U.S.A. 86, 4659-4663, 1989
A;Title: Isolation and DNA sequence of a cDNA clone encoding a lymphocyte adhesion recep
A;Reference number: A33935; MUID:89282830; PMID:2471974
A;Accession: A33935
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-362 <IDZ>
A;Cross-references: GB:M22452; NID:g176576; PID:g176577
C;Comment: This protein was isolated from the herpes papio induced B cell lymphoma.
C;Genetics:
A;Gene: CD44; ECMRIII; Hermes-1 antigen
C;Superfamily: human cell adhesion protein CD44
C;Keywords: cytoskeleton; extracellular matrix; glycoprotein; lipoprotein; surface antig
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-362/Product: lymphocyte adhesion receptor #status predicted <MAT>
F;21-270/Domain: extracellular #status predicted <EXT>
F;271-290/Domain: transmembrane #status predicted <TMM>
F;291-362/Domain: intracellular #status predicted <CYT>
F;25,57,100,110,120,256/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;296/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 13.4%; Score 222.5; DB 2; Length 362;
Best Local Similarity 25.5%; Pred. No. 2.4e-10;
Matches 92; Conservative 43; Mismatches 127; Indels 99; Gaps 16;

```
Qy         20 LVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQVETALK 79
              ||| ||   :: : ::||  ||  | |     :: ||| : |:    :|    |:| ||
Db         14 LVQLSL--AQIDLNITCRFEGIYHVEKNGRYSISRTEAADLCKAFNSTLPTMAQMEKALS 71

Qy         80 ASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDTWTNSC 139
                |||| ||:: :| ||| || ||  |  | || |     | |:  ||:|:|      |
Db         72 IGFETCRYGFI-EGHVVIPRIHPNSICAANNTGVYILTSNTS-QYDTYCFNASAPPGEDC 129

Qy        140 IPEIITTKDPIFNTQTATQTTEFIVSDST-YSVASPYSTIP------APT---------- 182
                 |:   : |      |   :  | | |     | | |      :||
Db        130 -----TSVTDLPNAFDGPITITIVNRDGTRYVKKGEYRTNPEDINPSSPTDDDVSSGSSS 184

Qy        183 ----------------TTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAF 226
                           |:|| |                    | ||:  |   :: ||
Db        185 ERSSTLGGYIFYNHFSTSPPIPDEDG---------------PWITDSTDRTPATRDQGAF 229

Qy        227 K----------NEAAGF------GG------------VPTALLVLALLFFGAAAGLGFCY 258
                       :|:||       ||            :|  |::|| |     |  | |
Db        230 DPSGGSHTTHGSESAGHSHGSREGGANTTSGPLRTPQIPEWLIILASL-LALALILAVC- 287

Qy        259 VKRYVKAPPFTNKNQQKEMIETK---VVKEEKANDSNPNEESKKTDK----NPEESKSPS 311
                  :          |:|:::       |:: |::  |  | ||  :     | | |::|
Db        288 ----IAVNSRRRCGQKKKLVINNGNGAVEDRKSSGLN-GEASKSQEMVHLVNKESSETPD 342

Qy        312 K 312
                :
Db        343 Q 343
```

RESULT 3
A34424
CD44 membrane glycoprotein precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 15-Jun-1990 #sequence_revision 15-Jun-1990 #text_change 01-Dec-2000
C;Accession: A34424; A34907
R;Nottenburg, C.; Rees, G.; St. John, T.
Proc. Natl. Acad. Sci. U.S.A. 86, 8521-8525, 1989
A;Title: Isolation of mouse CD44 cDNA: structural features are distinct from the primate
A;Reference number: A34424; MUID:90046829; PMID:2682651
A;Accession: A34424
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-365 <NOT>
A;Cross-references: GB:M27130; NID:g192530; PIDN:AAA37407.1; PID:g309161
R;Wolffe, E.J.; Gause, W.C.; Pelfrey, C.M.; Holland, S.M.; Steinberg, A.D.; August, J.T.
J. Biol. Chem. 265, 341-347, 1990
A;Title: The cDNA sequence of mouse Pgp-1 and homology to human CD44 cell surface antiger
A;Reference number: A34907; MUID:90094420; PMID:2403559
A;Accession: A34907
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 8-195,'G',197-365 <WOL>
A;Cross-references: GB:J05163; NID:g200334; PIDN:AAA39923.1; PID:g200335
C;Superfamily: human cell adhesion protein CD44
C;Keywords: cell adhesion; glycoprotein; membrane protein

Query Match 13.4%; Score 222.5; DB 2; Length 365;
Best Local Similarity 24.5%; Pred. No. 2.5e-10;
Matches 88; Conservative 56; Mismatches 150; Indels 65; Gaps 14;

```
Qy         15 WTTR---LLVQGSLR--AEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLA 69
              | |      |:| ||    ::: : |:||  |:  | |     :: ||| : |:    :|
Db          6 WHTAWGLCLLQLSLAHPHQQIDLNVTCRYAGVFHVEKNGRYSISRTEAADLCQAFNSTLP 65

Qy         70 GKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCY 129
                ||:: ||   |||| ||:: :| ||| || ||  |  | || |     :  :  ||:
Db         66 TMDQMKLALSKGFETCRYGFI-EGNVVIPRIHPNAICAANHTGVYILVTSNTSHYDTYCF 124

Qy        130 NSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDST-YSVASPYST---------- 177
              |:|    |     |:  : |:      |   :  | | ||    | |
Db        125 NASAPPEEDC-----TSVTDLPNSFDGPVTITIVNRDGTRYSKKGEYRTHQEDIDASNII 179

Qy        178 ----IPAPTTTPPAPAS----TSIPRRKKLICVTEVFMETSTMST-ETEPFVENKAAFK- 227
                       |     | |    | :|  :      : |   ||::| : :   ::: : :
Db        180 DDDVSSGSTIEKSTPESYILHTYLPTEQPTGDQDDSFFIRSTLATRDRDSSKDSRGSSRT 239

Qy        228 ----NEAAGFGG-----------------VPTALLVLALLFFGAAAGLGFCYVKRYVKA 265
                  :| ||                    :|  |::|| |     |  | |     :
Db        240 VTHGSELAGHSSANQDSGVTTTSGPMRRPQIPEWLIILASL-LALALILAVC-----IAV 293

Qy        266 FPFTNKNQQKEMI---ETKVVKEEKANDSNPNEESKKTDKNPEESKSPSKTTVRCLEAE 321
                    |:|:::       |:: | :: |  | ||  :     :| ||:|  :|: |:
Db        294 NSRRRCGQKKKLVINGGNGTVEDRKPSELN-GEASKSQEMVHLVNKEPSETPDQCMTAD 351
```

RESULT 4
B38745
cell adhesion molecule CD44 precursor, long form (meta-1) - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 24-Jan-1992 #sequence_revision 24-Jan-1992 #text_change 05-Nov-1999
C;Accession: B38745; A38745
R;Guenthert, U.; Hofmann, M.; Rudy, W.; Reber, S.; Zoeller, M.; HauBmann, I.; Matzku, S.
Cell 65, 13-24, 1991
A;Title: A new variant of glycoprotein CD44 confers metastatic potential to rat carcinoma
A;Reference number: A38745; MUID:91191552; PMID:1707342
A;Accession: B38745
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-503 <GUE>
A;Cross-references: GB:M61874; NID:g576534; PIDN:AAA53534.1; PID:g576535
A;Accession: A38745
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-223,386-503 <GU2>
A;Cross-references: GB:M61875
C;Keywords: cell adhesion

Query Match 13.4%; Score 222; DB 2; Length 503;
Best Local Similarity 25.7%; Pred. No. 4.1e-10;
Matches 84; Conservative 40; Mismatches 165; Indels 38; Gaps 11;

```
Qy         15 WTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQV 74
              |    |:| ||  ::: : ::||  |:  | |     :: ||| : |     :|    |:
Db         10 WGLLCLLQLSLAQQQIDLNITCRYAGVFHVEKNGRYSISRTEAADLCEAFNTTLPTMAQM 69

Qy         75 ETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDT 134
              | ||:  |||| ||:: :| ||| || ||  |  | || |     :  :  ||:|:|
Db         70 ELALRKGFETCRYGFI-EGHVVIPRIHPNAICAANNTGVYILLASNTSHYDTYCFNASAP 128
```

```
Qy        135 WTNSCIPEIITTKDPIFNTQTATQTTEFIVSDST-YSVASPYST----IPAPT------- 182
Db        129 LEEDC-----TSVTDLPNSFDGPVTITIVNRDGTRYSKKGEYRTHQEDIDASNIIDEDVS 183

Qy        183 -------TTPPA-PASTSIPRRKKLICVTEVFMETSTMST-ETEPFVENKAAFKNEAAGF 233
Db        184 SGSTIEKSTPEGYILHTDLPTSQPTGDRDDAFFIGSTLATIATTPWVSAHTKQNQERTQW 243

Qy        234 GGV---PTALLVLALLFFGAAAGLGFCYVKRYVK--AFPFTNKNQQKEMIETKVVKEEKA 288
Db        244 NPIHSNPEVLLQTTTRMTDIDRNSTSAHGENWTQEPQPPFNNHEYQDEE-ETPHATSTTW 302

Qy        289 NDSNPNEESKKTDK-----NPEESKSP 310
Db        303 ADPNSTTEEAATQKEKWFENEWQGKNP 329
```

RESULT 5
A35616
T-cell surface glycoprotein CD44 - hamster
C;Species: Cricetinae gen. sp. (hamster)
C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 21-Jul-2000
C;Accession: A35616
R;Aruffo, A.; Stamenkovic, I.; Melnick, M.; Underhill, C.B.; Seed, B.
Cell 61, 1303-1313, 1990
A;Title: CD44 is the principal cell surface receptor for hyaluronate.
A;Reference number: A35616; MUID:90304889; PMID:1694723
A;Accession: A35616
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-362 <ARU>
C;Superfamily: human cell adhesion protein CD44
C;Keywords: cell adhesion; glycoprotein; transmembrane protein

  Query Match 13.1%; Score 217.5; DB 2; Length 362;
  Best Local Similarity 25.4%; Pred. No. 6.2e-10;
  Matches 88; Conservative 46; Mismatches 149; Indels 63; Gaps 12;

```
Qy         24 SLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQVETALKASFE 83
Db         18 SLAHEQIDLNITCRYAGVFHVEKNGRYSISRTEAADLCQAFNSTLPTMDQMVMALSKGFE 77

Qy         84 TCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDTWTNSCIPEI 143
Db         78 TCRYGFI-EGHVVIPRIQPNAICAANHTGVYILTSNTS-HYDTYCFNASAPLEEDC---- 131

Qy        144 ITTKDPIFNTQTATQTTEFIVSDST-YSVASPYST----IPAPTTTPPAPAS-------- 190
Db        132 -TSVTDLPNSFEGPVTITIVNRDGTRYSKKGEYRTHQEDIDASNTTDDDVSSGSSSEKST 190

Qy        191 -------TSIPRRKKLICVTEVFMETSTMSTE--------------------TEPFVEN 222
Db        191 SGGYVFHTYLPTIHSTADQDDPYFIGSTMATRDQDSSMDPRGNSLTVTDGSKLTGHSSGN 250

Qy        223 KAAFKNEAAGFG---GVPTALLVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIE 279
Db        251 QDSGANTTSRPGRKPQIPEWLIVLASL-LALALILAVC-----IAVNSRRRCGQKKKLVI 304

Qy        280 TKVVKEEKANDSNPNEESKKTDKNPE----ESKSPSKTTVRCLEAE 321
Db        305 NS--GNGKVEDRKPSELNGEASKSQEMVHLVNKEPSETPDQFMTAD 348
```

RESULT 6
A53286
cell-surface glycoprotein CD44 precursor - bovine
N;Alternate names: CD44 protein
C;Species: Bos primigenius taurus (cattle)
C;Date: 02-May-1994 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C;Accession: A53286; S22123
R;Bosworth, B.T.; St John, T.; Gallatin, W.M.; Harp, J.A.
Mol. Immunol. 28, 1131-1135, 1991
A;Title: Sequence of the bovine CD44 cDNA: comparison with human and mouse sequences.
A;Reference number: A53286; MUID:92017904; PMID:1922105
A;Accession: A53286
A;Molecule type: mRNA
A;Residues: 1-366 <BOS>
A;Cross-references: EMBL:X62881; NID:g186; PIDN:CAA44675.1; PID:g187
A;Note: sequence extracted from NCBI backbone (NCBIN:63418, NCBIP:63419)
C;Superfamily: human cell adhesion protein CD44
C;Keywords: cell adhesion; glycoprotein; transmembrane protein
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-366/Product: cell-surface glycoprotein CD44 #status predicted <MAT>
F;274-294/Domain: transmembrane #status predicted <TMM>
F;25,57,100,110,120,222,260/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match 12.6%; Score 208.5; DB 2; Length 366;
  Best Local Similarity 23.8%; Pred. No. 3.3e-09;
  Matches 81; Conservative 44; Mismatches 142; Indels 73; Gaps 11;

```
Qy         20 LVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQVETALK 79
Db         14 LVQLSL--AQIDLNITCRYAGVFHVEKNGRYSISKTEAADLCKAFNSTLPTMAQMEAARN 71

Qy         80 ASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDTWTNSC 139
Db         72 IGFETCRYGFI-EGHVVIPRIHPNSICAANNTGVYILTSNTS-QYDTICFNASAPPGEDC 129

Qy        140 IPEIITTKDPIFNTQTATQTTEFIVSDST-YSVASPYSTIP--------APTTTPPAPAS 190
Db        130 -----TSVTDLPNAFEGPITITIVNRDGTRYTKKGEYRTNPEDINPSVVSPSSPPDDEMS 184

Qy        191 TSIPRRKKLICVTEVF----------------------METSTMSTETEPFVENKAAFK 227
Db        185 SGSPSERSTSGGYSIFHTHLPTVHPSRPRRPWSQRAENTSDTRDYGSSHDPSGRSYTTHA 244

Qy        228 NEAAGFGG-----------------VPTALLVLALLFFGAAAGLGFCYVKRYVKAFPFT 269
Db        245 SESAGHSSGSEEHGANTTSGPMRKPQIPEWLIILASL-LALALILAVC-----IAVNSRR 298

Qy        270 NKNQQKEMIETKVVKEEKANDSNPNEESKKTDKNPEESKS 309
Db        299 RCGQKKKLVIN--------NGNGTMEERKPSGLNGEASKS 330
```

RESULT 7
JH0417
cell adhesion molecule CD44 - human
C;Species: Homo sapiens (man)
C;Date: 23-Nov-1991 #sequence_revision 23-Nov-1991 #text_change 01-Dec-2000
C;Accession: JH0417; A32376; G02251; A32377
R;Harn, H.J.; Isola, N.; Cooper, D.L.
Biochem. Biophys. Res. Commun. 178, 1127-1134, 1991
A;Title: The multispecific cell adhesion molecule CD44 is represented in reticulocyte cDN
A;Reference number: JH0417; MUID:91337049; PMID:1840487
A;Accession: JH0417
A;Molecule type: mRNA
A;Residues: 1-361 <HAR>
A;Cross-references: GB:M59040; NID:g180129; PIDN:AAA51950.1; PID:g180130
A;Experimental source: reticulocyte
A;Note: the authors translated the codon ATG for residues 63, 66 and 239 as Trp and GGA
R;Stamenkovic, I.; Amiot, M.; Pesando, J.M.; Seed, B.
Cell 56, 1057-1062, 1989
A;Title: A lymphocyte molecule implicated in lymph node homing is a member of the cartil
A;Reference number: A32376; MUID:89168434; PMID:2466575
A;Accession: A32376
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-238,'E',240-361 <STA>
A;Cross-references: GB:M24915; NID:g180196; PIDN:AAA35674.1; PID:g180197
R;Bosch, P.P.; Stevens, J.W.; Buckwalter, J.A.; Midura, R.J.
submitted to the EMBL Data Library, November 1995
A;Reference number: H00921
A;Accession: G02251
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA

```
A;Residues: 1-25,'M',27-108,'S',110-361 <BOS>
A;Cross-references: EMBL:U40373; NID:g1101785; PID:g1101786
R;Goldstein, L.A.; Zhou, D.F.H.; Picker, L.J.; Minty, C.N.; Bargatze, R.F.; Ding, J.F.;
Cell 56, 1063-1072, 1989
A;Title: A human lymphocyte homing receptor, the hermes antigen, is related to cartilage
A;Reference number: A32377; MUID:89168435; PMID:2466576
A;Accession: A32377
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-108,'S',110-293,'S' <GOL>
A;Cross-references: GB:M25078; NID:g186660; PIDN:AAA36138.1; PID:g186661
C;Superfamily: human cell adhesion protein CD44
C;Keywords: alternative splicing; cell adhesion; surface antigen; transmembrane protein
F;269-285/Domain: transmembrane #status predicted <TMM>

  Query Match             12.4%;  Score 205;  DB 2;  Length 361;
  Best Local Similarity   24.4%;  Pred. No. 6.3e-09;
  Matches   83;  Conservative   46;  Mismatches  145;  Indels   66;  Gaps   14;

Qy         25 LRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQVETALKASFET 84
              |    :: : ::||  |:  | |      :: ||| : |:    :|    |:| ||   |||
Db         17 LSLAQIDLNITCRFAGVFHVEKNGRYSISRTEAADLCKAFNSTLPTMAQMEKALSIGFET 76

Qy         85 CSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDTWTNSCIPEII 144
              | ||:: :| ||| || ||  | |  || |     | |:  ||:|:|       |
Db         77 CRYGFI-EGHVVIPRIHPNSICAANNTGVYILTYNTS-QYDTYCFNASAPPEEDC----- 129

Qy        145 TTKDPIFNTQTATQTTEFIVSDST-YSVASPYSTIPA---PTTTPPAPASTSIPRRKKLI 200
              |:   : |     |   :  | | |     | | |     |:       |:     :
Db        130 TSVTDLPNAFDGPITITIVNRDGTRYVQKGEYRTNPEDIYPSNPTDDDVSSGSSSERSST 189

Qy        201 CVTEVFMETSTM-------------STETEPFVENKAAFK----------NEAAGF---- 233
                :|   ||:              ||:  |   ::  |          :|: |
Db        190 SGGYIFYTFSTVHPIPDEDSPWITDSTDRIPATRDQDTFHPSGGSHTTHGSESDGHSHGS 249

Qy        234 --GG------------VPTALLVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIE 279
                ||            :|  |::|| |    |  |  |     :          |:|:::
Db        250 QEGGANTTSGPIRTPQIPEWLIILASL-LALALILAVC-----IAVNSRRRCGQKKKLVI 303

Qy        280 TK---VVKEEKANDSNPNEESKKTDK----NPEESKSPSK 312
                  |:: | :  |  | ||  :     | | |::| :
Db        304 NSGNGAVEDRKPSGLN-GEASKSQEMVHLVNKESSETPDQ 342


RESULT 8
S45305
CD44 antigen precursor - dog (fragment)
C;Species: Canis lupus familiaris (dog)
C;Date: 20-Oct-1994 #sequence_revision 21-Jul-1995 #text_change 21-Jul-2000
C;Accession: S45305
R;Milde, K.F.; Alejandro, R.; Mintz, D.H.; Pastori, R.L.
Biochim. Biophys. Acta 1218, 112-114, 1994
A;Title: Molecular cloning of the canine CD44 antigen cDNA.
A;Reference number: S45305; MUID:94250687; PMID:7514890
A;Accession: S45305
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-351 <MIL>
A;Cross-references: EMBL:Z27115; NID:g473226; PIDN:CAA81630.1; PID:g473227
C;Superfamily: human cell adhesion protein CD44

  Query Match             12.2%;  Score 202.5;  DB 2;  Length 351;
  Best Local Similarity   23.4%;  Pred. No. 9.6e-09;
  Matches   84;  Conservative   49;  Mismatches  145;  Indels   81;  Gaps   15;

Qy         15 WTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQV 74
              |   ||     |   :: : ::||  |:  | |      :: | | : |:    :|     |:
Db          3 WGLCLL---RLSLAQIDLNITCRYAGVFHVEKNGRYSISRTAAADLCKAFNSTLPTMAQM 59

Qy         75 ETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDT 134
              | ||   |||| ||:: :| ||| || ||  |  |  || |  :  : |:  ||:|:|
Db         60 ERALSVGFETCRYGFI-EGHVVIPRIQPNAICAANHTGVYI-LISNTSQYDTYCFNASAP 117

Qy        135 WTNSCIPEIITTKDPIFNTQTATQTTEFIVSDST-YSVASPYSTIPA---PTTTPPAPAS 190
              |      |:   : |      |   :  | | ||    | | |     |:       |
Db        118 PEEDC-----TSVTHLPNAFDGPITITIVNRDGTRYSQKGEYRTNPEDINPSNPTDDDVS 172

Qy        191 TSIPRRKKLICVTEVF-----------METSTMSTETE-------------PFVENKAA 225
              :    :       :|            ::| :|: ::               |   :
Db        173 SGSSSERSTSAGYNIFHTHLPTAYPTEDQDSSRVSSNSDHTPITKDHDSSVHPSERSHTT 232

Qy        226 FKNEAAGF------GG------------VPTALLVLALLFFGAAAGLGFCYVKRYVKAFP 267
                :|:||       ||            :|  |::|| |    |  |  |
Db        233 HGSESAGHSSGSQEGGANTTSGPMRKPQIPEWLIILASL-LALALILAVCIA-------- 283

Qy        268 FTNKNQQKEMIETKVVKEEKAN----DSNP---NEESKKTDK-----NPEESKSPSKTT 314
              | ::   : | :       |   |   | |: |: :     | | |::| : |
Db        284 ---VNSRRRCGQKKKLVINNGNGAVGDRKPSGINGEASKSQEMVHLVNKEPSETPDQYT 339


RESULT 9
A47195
lymphocyte homing receptor isoform CD44 - human
C;Species: Homo sapiens (man)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 18-Nov-1994
C;Accession: A47195
R;Screaton, G.R.; Bell, M.V.; Jackson, D.G.; Cornelis, F.B.; Gerth, U.; Bell, J.I.
Proc. Natl. Acad. Sci. U.S.A. 89, 12160-12164, 1992
A;Title: Genomic structure of DNA encoding the lymphocyte homing receptor CD44 reveals at
A;Reference number: A47195; MUID:93101687; PMID:1465456
A;Accession: A47195
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-742 <SCR>
A;Note: sequence inconsistent with the nucleotide translation
A;Note: sequence extracted from NCBI backbone (NCBIN:120731, NCBIN:120737, NCBIN:120739,
, NCBIN:120764, NCBIN:120766, NCBIN:120770, NCBIN:120772, NCBIN:120774, NCBIN:120776, NCE

  Query Match             12.2%;  Score 202;  DB 2;  Length 742;
  Best Local Similarity   26.0%;  Pred. No. 2.7e-08;
  Matches   61;  Conservative   30;  Mismatches   84;  Indels   60;  Gaps    7;

Qy         25 LRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQVETALKASFET 84
              |    :: : ::||  |:' | |      :: ||| : |:    :|    |:| ||   |||
Db         17 LSLAQIDLNITCRFAGVFHVEKNGRYSISRTEAADLCKAFNSTLPTMAQMEKALSIGFET 76

Qy         85 CSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDTWTNSCI---- 140
              | ||:: :| ||| || ||  |  || |     | |:  ||:|:|       |
Db         77 CRYGFI-EGHVVIPRIHPNSICAANNTGVYILTSNTS-QYDTYCFNASAPPEEDCTSVTD 134

Qy        141 ------------------------------PEIITTKDPIFNTQTATQTTEFIVSDSTYS 170
                                           || |   :|  :  ::  ::|    :   |
Db        135 LPNAFDGPITITIVNRDGTRYVQKGEYRTNPEDIYPSNPTDDDVSSGSSSERSSTSGGY- 193

Qy        171 VASPYSTI-PAPTTTPP--------APASTSIPRRKKLICVTEVFMETSTMSTET 216
                :   :||: | |    |          ||:|               | ||  :|||
Db        194 IFYTFSTVHPIPDEDSPWITDSTDRIPATT-------------LMSTSATATET 234


RESULT 10
JH0518
lymphocyte homing receptor CD44, splice form CD44R1 - human
N;Alternate names: cell adhesion molecule core protein CD44E, keratinocyte; cell surface
N;Contains: lymphocyte homing receptor CD44, splice form CD44R1; lymphocyte homing recept
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 18-Aug-2000
C;Accession: JH0518; JH0519; PH0859; A39209; A42402; C42402; A53029; S16147
R;Dougherty, G.J.; Lansdorp, P.M.; Cooper, D.L.; Humphries, R.K.
J. Exp. Med. 174, 1-5, 1991
A;Title: Molecular cloning of CD44R1 and CD44R2, two novel isoforms of the human CD44 lym
A;Reference number: JH0518; MUID:91277598; PMID:2056274
A;Accession: JH0518
A;Molecule type: mRNA
A;Residues: 1-426 <DOU>
A;Experimental source: lymphocytes, cell line KG1a
A;Accession: JH0519
```

```
A;Molecule type: mRNA
A;Residues: 1-223,288-426 <DO2>
A;Experimental source: lymphocyte, cell line KG1a
R;Cooper, D.L.; Dougherty, G.; Harn, H.J.; Jackson, S.; Baptist, E.W.; Byers, J.; Datta,
Biochem. Biophys. Res. Commun. 182, 569-578, 1992
A;Title: The complex CD44 transcriptional unit: alternative splicing of three internal e
A;Reference number: PH0859; MUID:92134271; PMID:1734871
A;Accession: PH0859
A;Molecule type: DNA
A;Residues: 223-357 <COO>
R;Brown, T.A.; Bouchard, T.; St. John, T.; Wayner, E.; Carter, W.G.
J. Cell Biol. 113, 207-221, 1991
A;Title: Human keratinocytes express a new CD44 core protein (CD44E) as a heparan-sulfat
A;Reference number: A39209; MUID:91177958; PMID:2007624
A;Accession: A39209
A;Molecule type: mRNA
A;Residues: 184-376 <BRO>
A;Cross-references: GB:X55938; NID:g29802; PIDN:CAA39404.1; PID:g930047
R;Jackson, D.G.; Buckley, J.; Bell, J.I.
J. Biol. Chem. 267, 4732-4739, 1992
A;Title: Multiple variants of the human lymphocyte homing receptor CD44 generated by ins
A;Reference number: A42402; MUID:92165834; PMID:1537855
A;Accession: A42402
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 217-223,288-359 <JAC>
A;Note: sequence extracted from NCBI backbone (NCBIN:83964, NCBIP:83965)
A;Note: variant B
A;Accession: C42402
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 217-320 <JA2>
A;Note: sequence extracted from NCBI backbone (NCBIN:83968, NCBIP:83969)
A;Note: variant D
R;Shepley, M.P.; Racaniello, V.R.
J. Virol. 68, 1301-1308, 1994
A;Title: A monoclonal antibody that blocks poliovirus attachment recognizes the lymphocy
A;Reference number: A53029; MUID:94149816; PMID:7508992
A;Accession: A53029
A;Status: preliminary
A;Molecule type: protein
A;Residues: 67-76,'X',78-89 <SHE>
C;Genetics:
A;Gene: GDB:CD44; MDU2; MDU3; MI
A;Cross-references: GDB:120739; OMIM:107269
A;Map position: 11pter-11p13
A;Introns: 35/1; 65/1; 133/1
C;Superfamily: human cell adhesion protein CD44
C;Keywords: alternative splicing; cell adhesion; chondroitin sulfate proteoglycan; glyco
F;1-426/Product: lymphocyte homing receptor CD44, splice form CD44R1 #status predicted <
F;1-223,288-426/Product: lymphocyte homing receptor CD44, splice form CD44R2 #status pre
F;299/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;354/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted

  Query Match             12.0%;  Score 199.5;  DB 2;  Length 426;
  Best Local Similarity   22.2%;  Pred. No. 2.1e-08;
  Matches   83;  Conservative   48;  Mismatches  138;  Indels  105;  Gaps   12;

Qy         25 LRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQVETALKASFET 84
              |   :: : ::||  |:  | |     :: ||| : |:    :|    |:| ||   |||
Db         17 LSLAQIDLNITCRFAGVFHVEKNGRYSISRTEAADLCKAFNSTLPTMAQMEKALSIGFET 76

Qy         85 CSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDTWTNSCI---- 140
              | ||:: :| ||| || ||  |  |  || |     | |:  ||:|:|      |
Db         77 CRYGFI-EGHVVIPRIHPNSICAANNTGVYILTSNTS-QYDTYCFNASAPPEEDCTSVTD 134

Qy        141 ------------------------------PEIITTKDPIFNTQTATQTTE-------FI 163
                                           || |   :|  :  ::  ::|       :|
Db        135 LPNAFDGPITITIVNRDGTRYVQKGEYRTNPEDIYPSNPTDDDVSSGSSSERSSTSGGYI 194

Qy        164 -----------------VSDST-----YSVASPYSTIPAPTTTPPAPASTSIPRRKKLIC 201
                             ::|||      :: | :||   ||  |       : |   |
Db        195 FYTFSTVHPIPDEDSPWITDSTDRIPATNMDSSHSTTLQPTANPNTGLVEDLDRTGPLSM 254

Qy        202 VTEVFMETSTMSTETEPFVENK---------AAFKNEAAGFGGVP------TALLVLALL 246
              |:    : : ||  |   |:|         :: :|:  |       |      | ||
Db        255 TTQQ-SNSQSFSTSHEGLEEDKDHPTTSTLTSSNRNDVTGGRRDPNHSEGSTTLL----- 308

Qy        247 FFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKV-----VKEEKANDSNPN-EESKKT 300
                          : |   :| | :::    : :       |     ||| |   |
Db        309 ------------EGYTSHYPHTKESRTFIPVTSAKTGSFGVTAVTVGDSNSNVNRSLSG 355

Qy        301 DKNPEESKSPSKTT 314
              |::       | ||
Db        356 DQDTFHPSGGSHTT 369


RESULT 11
S13530
CD44E protein, epithelial - human
C;Species: Homo sapiens (man)
C;Date: 18-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 21-Jul-2000
C;Accession: S13530
R;Stamenkovic, I.; Aruffo, A.; Amiot, M.; Seed, B.
EMBO J. 10, 343-348, 1991
A;Title: The hematopoietic and epithelial forms of CD44 are distinct polypeptides with d
A;Reference number: S13530; MUID:91122041; PMID:1991450
A;Accession: S13530
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-493 <STA>
A;Cross-references: EMBL:X55150; NID:g29800; PIDN:CAA38951.1; PID:g29801
C;Keywords: transmembrane protein

  Query Match             12.0%;  Score 199.5;  DB 2;  Length 493;
  Best Local Similarity   23.6%;  Pred. No. 2.6e-08;
  Matches   66;  Conservative   39;  Mismatches  100;  Indels   75;  Gaps    8;

Qy         25 LRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQVETALKASFET 84
              |   :: : ::||  |:  | |     :: ||| : |:    :|    |:| ||   |||
Db         17 LSLAQIDLNITCRFAGVFHVEKNGRYSISRTEAADLCKAFNSTLPTMAQMEKALSIGFET 76

Qy         85 CSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDTWTNSCI---- 140
              | ||:: :| ||| || ||  |  |  || |     | |:  ||:|:|      |
Db         77 CRYGFI-EGHVVIPRIHPNSICAANNTGVYILTYNTS-QYDTYCFNASAPPEEDCTSVTD 134

Qy        141 ------------------------------PEIITTKDPIFNTQTATQTTE-------FI 163
                                           || |   :|  :  ::  ::|       :|
Db        135 LPNAFDGPITITIVNRDGTRYVQKGEYRTNPEDIYPSNPTDDDVSSGSSSERSSTSGGYI 194

Qy        164 -----------------VSDST-----YSVASPYSTIPAPTTTPPAPASTSIPRRKKLIC 201
                             ::|||      :: | :||   ||  |       : |   |
Db        195 FYTFSTVHPIPDEDSPWITDSTDRIPRTNMDSSHSTTLQPTANPNTGLVEDLDRTGPLSM 254

Qy        202 VTEVFMETSTMSTETEPFVENK---------AAFKNEAAG 232
              |:    : : ||  |   |:|         :: :|:  |
Db        255 TTQQ-SNSQSFSTSHEGLEEDKDHPTTSTLTSSNRNDVTG 293


RESULT 12
I77371
CD44R5 - human
C;Species: Homo sapiens (man)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jul-2000
C;Accession: I77371
R;Tanabe, K.K.; Nishi, T.; Saya, H.
Mol. Carcinog. 7, 212-220, 1993
A;Title: Novel variants of CD44 arising from alternative splicing: changes in the CD44 a
A;Reference number: I57483; MUID:93356912; PMID:8352881
A;Accession: I77371
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-395 <RES>
A;Cross-references: GB:S66400; NID:g435697; PIDN:AAB27919.1; PID:g435700
C;Genetics:
A;Gene: GDB:CD44
```

A;Cross-references: GDB:120739; OMIM:107269
A;Map position: 11pter-11p13
A;Introns: 257/1
C;Superfamily: human cell adhesion protein CD44

Query Match 11.9%; Score 198; DB 2; Length 395;
Best Local Similarity 23.0%; Pred. No. 2.6e-08;
Matches 85; Conservative 49; Mismatches 145; Indels 90; Gaps 14;

```
Qy         25 LRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQVETALKASFET 84
Db         17 LSLAQIDLNITCRFAGVFHVEKNGRYSISRTEAADLCKAFNSTLPTMAQMEKALSIGFET 76

Qy         85 CSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDTWTNSCI---- 140
Db         77 CRYGFI-EGHVVIPRIHPNSICAANNTGVYILTSNTS-QYDTYCFNASAPPEEDCTSVTD 134

Qy        141 ----------------------------PEIITTKDPIFNTQTATQTTEFIVSDS--- 167
Db        135 LPNAFDGPITITIVNRDGTRYVQKGEYRTNPEDIYPSNPTDDDVSSGSSSERSSTSGGYI 194

Qy        168 --TYSVASPY----------STIPAPTTTPPAPASTSI-PRRKKLICVTEVFMETSTMST 214
Db        195 FYTFSTVHPIPDEDSPWITDSTDRIPATNMDSSHSTTLQPTANPNTGLVEDLDRTGPLSM 254

Qy        215 ET------EPFVENKAAFKNEAAGF------GG------------VPTALLVLALLFFGA 250
Db        255 TTRDQDTFHPSGGSHTTHGSESDGHSHGSQEGGANTTSGPIRTPQIPEWLIILASL-LAL 313

Qy        251 AAGLGFCYVKRYVKAFPFTNKNQQKEMIETK---VVKEEKANDSNPNEESKKTDK----N 303
Db        314 ALILAVC-----IAVNSRRRCGQKKKLVINSGNGAVEDRKPSGLN-GEASKSQEMVHLVN 367

Qy        304 PEESKSPSK 312
Db        368 KESSETPDQ 376
```

RESULT 13
I37369
epican - human
C;Species: Homo sapiens (man)
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 03-Aug-2001
C;Accession: I37369; S24631
R;Kugelman, L.C.; Ganguly, S.; Haggerty, J.G.; Weissman, S.M.; Milstone, L.M.
J. Invest. Dermatol. 99, 866-891, 1992
A;Title: The core protein of epican, a heparan sulfate proteoglycan on keratinocytes, is
A;Reference number: I37369; PMID:1281868
A;Note: corrected and republished from J. Invest. Dermatol. 99, 381-385, 1992
A;Accession: I37369
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-699 <RES>
A;Cross-references: EMBL:X66733; NID:g31190; PID:g31191

Query Match 11.9%; Score 196.5; DB 2; Length 699;
Best Local Similarity 25.4%; Pred. No. 6.9e-08;
Matches 60; Conservative 34; Mismatches 93; Indels 49; Gaps 7;

```
Qy         25 LRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQVETALKASFET 84
Db         17 LSLAQIDLNITCRFAGVFHVEKNGRYSISRTEAADLCKAFNSTLPTMAQMEKALSIGFET 76

Qy         85 CSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDTWTNSCI---- 140
Db         77 CRYGFI-EGHVVIPRIHPNSICAANNTGVYILTSNTS-QYDTYCFNASAPPEEDCTSVTD 134

Qy        141 ----------------------------PEIITTKDPIFNTQTATQTTEFIVSDSTYS 170
Db        135 LPNAFDGPITITIVNRDGTRYVQKGEYRTNPEDIYPSNPTDDDVSSGSSSERSSTSGGY- 193

Qy        171 VASPYSTI-PAPTTTPP--APASTSIPRRKKLICVTEVFMETSTMSTETEPFVENK 223
```

Db 194 IFYTFSTVHPIPDEDSPWITDSTDRIPATS---------TSSNTISAGWEPNEENE 240

RESULT 14
S24240
lymphocyte surface antigen CD44 precursor - horse
C;Species: Equus caballus (domestic horse)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 21-Jul-2000
C;Accession: I46245; S24240
R;Tavernor, A.S.; Deverson, E.V.; Coadwell, W.J.; Lunn, D.P.; Zhang, C.; Davis, W.; Butch
Immunogenetics 37, 474-477, 1993
A;Title: Molecular cloning of equine CD44 cDNA by a COS cell expression system.
A;Reference number: I46245; MUID:93170897; PMID:8436424
A;Accession: I46245
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-359 <TA2>
A;Cross-references: EMBL:X66862; NID:g1059; PIDN:CAA47331.1; PID:g1060
C;Superfamily: human cell adhesion protein CD44
C;Keywords: surface antigen; transmembrane protein

Query Match 11.2%; Score 186; DB 2; Length 359;
Best Local Similarity 24.6%; Pred. No. 2.1e-07;
Matches 84; Conservative 42; Mismatches 159; Indels 56; Gaps 11;

```
Qy         25 LRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQVETALKASFET 84
Db         17 LSLAQIDLNITCRYAGVFHVEKNGRYSISRTEAADLCKAFNSTLPTMAQMQKALNIGFET 76

Qy         85 CSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDTWTNSCI---- 140
Db         77 CRIGFI-EGHVVIPPIHPNSICAANNTGVYILTSNTS-QYDTYCFNASAPPEEDCTSVTD 134

Qy        141 -------PEIIT----------------TKDPIFNTQTATQTTEFIVSDSTYSVASPYST 177
Db        135 LPNAFEGPITITIVNRDGTRYTKKGEYRTNPEDINPSTPADDDVSSGSSSERSTSGGYSI 194

Qy        178 IPA--PTTTP------------PAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENK 223
Db        195 FHTHLPTTRPTQDQSSPWVSDSPEKTPTTKDRASGGRAQTTHGSETSGHSTGSQ---EGG 251

Qy        224 AAFKNEAAGFGGVPTALLVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETK-- 281
Db        252 ASTTSGPIRRPQIPEWLIILASL-LALALILAVC-----IAVNSRRRCGQKKKLVINNGN 305

Qy        282 -VVKEEKANDSNPNEESKKTDKNPEESKSPSKTTVRCLEAE 321
Db        306 GAVDDRKASGLN-GEASRSQEMVHLVNKESSETQDQFMTAD 345
```

RESULT 15
I50421
aggrecan precursor - chicken
N;Alternate names: cartilage chondroitin sulfate proteoglycan core protein
C;Species: Gallus gallus (chicken)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jan-2000
C;Accession: I50421; S39796; S27356; A25442; A32002; I50216; A37072; B37072
R;Li, H.; Schwartz, N.B.; Vertel, B.M.
J. Biol. Chem. 268, 23504-23511, 1993
A;Title: cDNA cloning of chick cartilage chondroitin sulfate (aggrecan) core protein and
A;Reference number: A48884; MUID:94043149; PMID:8226878
A;Accession: I50421
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-2109 <LIX>
A;Cross-references: GB:L21913; NID:g416133; PIDN:AAB19128.1; PID:g416134
R;Chandrasekaran, L.; Tanzer, M.L.
Biochem. J. 296, 885-887, 1993
A;Reference number: S39796; MUID:94107258; PMID:8280087
A;Contents: annotation; erratum
A;Accession: S39796
A;Molecule type: mRNA
A;Residues: 1-361,'DL',364-600,'R',602-999,'R',1001-1028,'P',1030-1250,'D',1252-1602,'A'

```
A;Cross-references: GB:M88101
R;Chandrasekaran, L.; Tanzer, M.L.
Biochem. J. 288, 903-910, 1992
A;Title: Molecular cloning of chicken aggrecan. Structural analyses.
A;Reference number: S27356; MUID:93111968; PMID:1339285
A;Accession: S27356
A;Molecule type: mRNA
A;Residues: 1-361,'DL',364-600,'R',602-999,'R',1001-1028,'P',1030-1250,'D',1252-1549,'T'
A;Cross-references: EMBL:M88101
R;Sai, S.; Tanaka, T.; Kosher, R.A.; Tanzer, M.L.
Proc. Natl. Acad. Sci. U.S.A. 83, 5081-5085, 1986
A;Title: Cloning and sequence analysis of a partial cDNA for chicken cartilage proteogly
A;Reference number: A25442; MUID:86259736; PMID:3460082
A;Accession: A25442
A;Molecule type: mRNA
A;Residues: 1693-1795,'G',1797-1855,1894-2109 <SAI>
A;Cross-references: GB:M13993; NID:g211654; PIDN:AAA48720.1; PID:g211655
A;Experimental source: sternal cartilage
R;Tanaka, T.; Har-El, R.; Tanzer, M.L.
J. Biol. Chem. 263, 15831-15835, 1988
A;Title: Partial structure of the gene for chicken cartilage proteoglycan core protein.
A;Reference number: A32002; MUID:89008500; PMID:3170613
A;Accession: A32002
A;Molecule type: DNA
A;Residues: 1893-1987,'S',1989-2022 <TAN>
A;Note: the authors translated the codon TCC for residue 1787 as Phe
R;Krueger, R.C.
J. Biol. Chem. 265, 12088-12097, 1990
A;Title: Chick cartilage chondroitin sulfate proteoglycan core protein: II. Nucleotide s
A;Reference number: I50216; MUID:90307744; PMID:1694853
A;Accession: I50216
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 'PA',1044-1559 <KRU>
A;Cross-references: GB:M38187; NID:g211685; PIDN:AAA48731.1; PID:g555441
R;Krueger Jr., R.C.; Fields, T.A.; Hildreth IV, J.; Schwartz, N.B.
J. Biol. Chem. 265, 12075-12087, 1990
A;Title: Chick cartilage chondroitin sulfate proteoglycan core protein. I. Generation an
A;Reference number: A37072; MUID:90307743; PMID:2365711
A;Accession: A37072
A;Molecule type: protein
A;Residues: 998-1015,'X',1017-1019,'X',1021-1023 <KR2>
A;Note: amino end of 86K core peptide CS-A
A;Accession: B37072
A;Molecule type: protein
A;Residues: 1247-1250,'D',1252-1272,'X',1274-1275 <KR3>
A;Note: amino end of 75K core peptide CS-B
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EG
C;Keywords: alternative splicing
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-2109/Product: aggrecan #status predicted <MAT>
F;44-131/Domain: immunoglobulin homology <IMM>
F;166-243/Domain: link protein repeat homology <LNK1>
F;264-346/Domain: link protein repeat homology <LNK2>
F;537-614/Domain: link protein repeat homology <LNK3>
F;635-716/Domain: link protein repeat homology <LNK4>
F;1859-1890/Domain: EGF homology <EGF>
F;1897-2017/Domain: C-type lectin homology <LCH>
F;2024-2080/Domain: complement factor H repeat homology <FHD>

  Query Match             9.0%;  Score 148.5;  DB 1;  Length 2109;
  Best Local Similarity   27.0%;  Pred. No. 0.002;
  Matches   74;  Conservative   27;  Mismatches  104;  Indels   69;  Gaps   13;

Qy          50 QQLNFTEAKEACRLLGLSLAGKDQVETALKASFETCSYGWVGDGFV--VISRISPNPKCG 107
               ::  | || : |  ||  ||   ::  | |   : || ||: |  |   |||  ||  ||
Db         258 EKFTFQEAFDKCHSLGARLATTGELYLAWKDGMDMCSAGWLADRSVRYPISRARPN--CG 315

Qy         108 KNGVGV-LIWKVPV--------SRQFAAYCYNSSDTWTNSCIP-----EIITTKDPIFNT 153
                 | |||  ::  |        | :: | ||:  |    : :|     |: |     |
Db         316 GNLVGVRTVYLNPANQTGYPHPSSRYDAICYSGDD--FEALVPGLFTDEVGTELGSAFTI 373

Qy         154 QTATQT----------TEFIVSDSTYSVASPYSTIPAPTTTPPAPASTSIPRRKKLICVT 203
               ||  |||          ||    ::  |:|:        | |    | | :|     :  |
Db         374 QTVTQTEVELPLPRNVTE---EEARGSIATLEPMEITATATELYEAFTVLPD----LFAT 426

Qy         204 EVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTALLVLALLFFGAAAGLGFCYVKRYV 263
               | :||::       |  ||    : |  |   ||                        |  |
Db         427 SVTVETAS------PREEN--VTREEITGIWAVPEE------------------VTTSV 459

Qy         264 KAFPFTNKNQQKEMIETKVVKEEKANDSNPNEES 297
                 ||       | |   |:|  |  : |  ||
Db         460 SGTAFTT-----GMAEVSSVEEAIAVTATPGLES 488


Search completed: August 11, 2004, 11:54:36
Job time : 18 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 11, 2004, 11:43:31 ; Search time 13 Seconds
(without alignments)
1289.738 Million cell updates/sec

Title: US-10-063-510-6
Perfect score: 1657
Sequence: 1 MARCFSLVLLLTSIWTTRLL.....NPEESKSPSKTTVRCLEAEV 322

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	233.5	14.1	362	1	CD44_CRIGR	P20944 cricetulus
2	222.5	13.4	362	1	CD44_PAPHA	P14745 papio hamad
3	222	13.4	503	1	CD44_RAT	P26051 rattus norv
4	212.5	12.8	431	1	CD44_MESAU	Q60522 m cd44 anti
5	210	12.7	778	1	CD44_MOUSE	P15379 mus musculu
6	208.5	12.6	366	1	CD44_BOVIN	Q29423 bos taurus
7	202.5	12.2	351	1	CD44_CANFA	Q28284 canis famil
8	202	12.2	742	1	CD44_HUMAN	P16070 h cd44 anti
9	186	11.2	359	1	CD44_HORSE	Q05078 equus cabal
10	150	9.1	537	1	PGCA_PIG	Q29011 sus scrofa
11	148.5	9.0	2109	1	PGCA_CHICK	P07898 gallus gall
12	145.5	8.8	3562	1	PGCV_CHICK	Q90953 gallus gall
13	145	8.8	277	1	TSG6_HUMAN	P98066 homo sapien
14	144	8.7	275	1	TSG6_MOUSE	O08859 mus musculu
15	143	8.6	2364	1	PGCA_BOVIN	P13608 bos taurus
16	139.5	8.4	2333	1	PGCA_CANFA	Q28343 canis famil
17	137	8.3	276	1	TSG6_RABIT	P98065 oryctolagus
18	131.5	7.9	2124	1	PGCA_RAT	P07897 rattus norv
19	131	7.9	2132	1	PGCA_MOUSE	Q61282 mus musculu
20	130.5	7.9	883	1	PGCB_MOUSE	Q61361 mus musculu
21	130.5	7.9	883	1	PGCB_RAT	P55068 rattus norv
22	130.5	7.9	2415	1	PGCA_HUMAN	P16112 homo sapien
23	128.5	7.8	912	1	PGCB_BOVIN	Q28062 bos taurus
24	126	7.6	3358	1	PGCV_MOUSE	Q62059 mus musculu
25	122	7.4	394	1	PGCA_RABIT	Q28670 oryctolagus
26	118.5	7.2	3396	1	PGCV_HUMAN	P13611 homo sapien
27	118	7.1	341	1	BRA1_MOUSE	Q9esm3 mus musculu
28	117	7.1	340	1	BRA1_HUMAN	Q9gzv7 homo sapien
29	116	7.0	2738	1	PGCV_RAT	Q9erb4 rattus norv
30	115	6.9	3381	1	PGCV_BOVIN	P81282 bos taurus
31	114.5	6.9	1257	1	PGCN_RAT	P55067 rattus norv
32	113	6.8	341	1	BRA1_RAT	Q9esm2 rattus norv
33	113	6.8	892	1	LDL2_XENLA	Q99088 xenopus lae
34	112.5	6.8	417	1	PGCB_FELCA	P41725 felis silve
35	112	6.8	390	1	VGLI_HSV11	P06487 herpes simp
36	111.5	6.7	402	1	ODP2_MYCPN	P75392 mycoplasma
37	109	6.6	1321	1	PGCN_HUMAN	O14594 homo sapien
38	108	6.5	354	1	PLK_HUMAN	P10915 homo sapien
39	107	6.5	354	1	PLK_RAT	P03994 rattus norv
40	107	6.5	355	1	PLK_CHICK	P07354 gallus gall
41	107	6.5	356	1	PLK_MOUSE	Q9qup5 mus musculu
42	107	6.5	1268	1	PGCN_MOUSE	P55066 mus musculu
43	105	6.3	354	1	PLK_BOVIN	P55252 bos taurus
44	104.5	6.3	862	1	PGCV_MACNE	Q28858 macaca neme
45	104	6.3	354	1	PLK_HORSE	Q28381 equus cabal

ALIGNMENTS

RESULT 1
CD44_CRIGR
ID CD44_CRIGR STANDARD; PRT; 362 AA.
AC P20944;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE CD44 antigen precursor (Phagocytic glycoprotein I) (PGP-1) (HUTCH-I)
DE (Extracellular matrix receptor-III) (ECMR-III) (GP90 lymphocyte
DE homing/adhesion receptor) (Hermes antigen) (Hyaluronate receptor).
GN CD44.
OS Cricetulus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90304889; PubMed=1694723;
RA Aruffo A., Stamenkovic I., Melnick M., Underhill C.B., Seed B.;
RT "CD44 is the principal cell surface receptor for hyaluronate.";
RL Cell 61:1303-1313(1990).
CC -!- FUNCTION: Main cell surface receptor for hyaluronate. Adhesion to
CC mucosal high endothelial venule and to types I and VI collagen.
CC Probably involved in matrix adhesion, lymphocyte activation and
CC lymph node homing.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- INDUCTION: By EBV.
CC -!- PTM: Extensively modified including N- and O-linked glycosylation,
CC addition of the glycosaminoglycan chondroitin sulfate, of sulfate,
CC of phosphate to cytoplasmic domain serine residues.
CC -!- SIMILARITY: Contains 1 link domain.
CC --

CC --
DR EMBL; M33827; AAA36967.1; -.
DR PIR; A35616; A35616.
DR HSSP; P98066; 1TSG.
DR InterPro; IPR001231; CD44_antigen.
DR InterPro; IPR000538; Link.
DR Pfam; PF00193; Xlink; 1.
DR PRINTS; PR00658; CD44.
DR PRINTS; PR01265; LINKMODULE.
DR ProDom; PD000918; Link; 1.
DR SMART; SM00445; LINK; 1.
DR PROSITE; PS01241; LINK; 1.
KW Cell adhesion; Transmembrane; Glycoprotein; Phosphorylation; Receptor;
KW Proteoglycan; Signal; Alternative splicing;
KW Pyrrolidone carboxylic acid.
FT SIGNAL 1 22 BY SIMILARITY.

```
FT   CHAIN        23    362       CD44 ANTIGEN.
FT   DOMAIN       23    269       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    270    290       POTENTIAL.
FT   DOMAIN       291   362       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN        50    121      LINK.
FT   DOMAIN       152   160       ARG/LYS-RICH (BASIC).
FT   DOMAIN       226   269       STEM.
FT   MOD_RES       23     23      PYRROLIDONE CARBOXYLIC ACID (BY
FT                                SIMILARITY).
FT   DISULFID      55    120      BY SIMILARITY.
FT   DISULFID      79     99      BY SIMILARITY.
FT   MOD_RES      292    292      PHOSPHORYLATION (BY PKC) (PARTIAL) (BY
FT                                SIMILARITY).
FT   MOD_RES      326    326      PHOSPHORYLATION (PARTIAL) (BY
FT                                SIMILARITY).
FT   CARBOHYD      27     27      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      59     59      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     102    102      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     112    112      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     122    122      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     174    174      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     256    256      N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   362 AA;  39775 MW;  E89EB434E9EEC948 CRC64;

  Query Match             14.1%;  Score 233.5;  DB 1;  Length 362;
  Best Local Similarity   26.3%;  Pred. No. 9.8e-12;
  Matches   91;  Conservative   45;  Mismatches  147;  Indels   63;  Gaps   12;

Qy          24 SLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQVETALKASFE 83
               ||  |:: : ::||  |:  | |     :: ||| : |:    :|   ||:  ||   ||
Db          18 SLAHEQIDLNITCRYAGVFHVEKNGRYSISRTEAADLCQAFNSTLPTMDQMVMALSKGFE 77

Qy          84 TCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDTWTNSCIPEI 143
               || ||:: :| ||| || ||  |  |  || |     |  :  ||:|:|      |
Db          78 TCRYGFI-EGHVVIPRIQPNAICAANHTGVYILTSNTS-HYDTYCFNASAPLEEDC---- 131

Qy         144 ITTKDPIFNTQTATQTTEFIVSDST-YSVASPYST----IPAPTTTPPAPAS-------- 190
                |:   : |:     |   :  | | ||    | |    | | ||    :|
Db         132 -TSVTDLPNSFEGPVTITIVNRDGTRYSKKGEYRTHQEDIDASNTTDDDVSSGSSSEKST 190

Qy         191 -------TSIPRRKKLICVTEVFMETSTMSTE--------------------TEPFVEN 222
                    | :|          : :   |||:|                     ||    |
Db         191 SGGYVFHTYLPTIHSTADQDDPYFIGSTMATRDQDSSMDPRGNSLTVTDGSKLTEHSSGN 250

Qy         223 KAAFKNEAAGFGG---VPTALLVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIE 279
               : :  |  :  ||   ||  |:||| |    |  |  |     :           |:|:::
Db         251 QDSGLNSTSRPGGKPRVPEWLIVLASL-LALALILAVC-----IAVNSRRRCGQKKKLVI 304

Qy         280 TKVVKEEKANDSNPNEESKKTDKNPE----ESKSPSKTTVRCLEAE 321
                    |  |  |:| : :  |: |     :| ||:|  : : |:
Db         305 NS--GNGKVEDRKPSELNGEASKSQEMVHLVNKEPSETPDQFMTAD 348


RESULT 2
CD44_PAPHA
ID   CD44_PAPHA     STANDARD;      PRT;   362 AA.
AC   P14745;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   CD44 antigen precursor (Phagocytic glycoprotein I) (PGP-1) (HUTCH-I)
DE   (Extracellular matrix receptor-III) (ECMR-III) (GP90 lymphocyte
DE   homing/adhesion receptor) (Hermes antigen) (Hyaluronate receptor).
GN   CD44.
OS   Papio hamadryas (Hamadryas baboon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Papio.
OX   NCBI_TaxID=9557;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 66-74.
RX   MEDLINE=89282830; PubMed=2471974;
RA   Idzerda R.L., Carter W.G., Nottenburg C., Wayner E.A.,
RA   Gallatin W.M., St John T.;
RT   "Isolation and DNA sequence of a cDNA clone encoding a lymphocyte
RT   adhesion receptor for high endothelium.";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:4659-4663(1989).
CC   -!- FUNCTION: Main cell surface receptor for hyaluronate. Adhesion to
CC       mucosal high endothelial venule and to types I and VI collagen.
CC       Probably involved in matrix adhesion, lymphocyte activation and
CC       lymph node homing.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- PTM: Extensively modified including N- and O-linked glycosylation,
CC       addition of the glycosaminoglycan chondroitin sulfate, of sulfate,
CC       of phosphate to cytoplasmic domain serine residues.
CC   -!- SIMILARITY: Contains 1 link domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M22452; AAA35385.1; -.
DR   HSSP; P98066; 1TSG.
DR   InterPro; IPR001231; CD44_antigen.
DR   InterPro; IPR000538; Link.
DR   Pfam; PF00193; Xlink; 1.
DR   PRINTS; PR00658; CD44.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 1.
DR   SMART; SM00445; LINK; 1.
DR   PROSITE; PS01241; LINK; 1.
KW   Cell adhesion; Transmembrane; Glycoprotein; Phosphorylation; Receptor;
KW   Proteoglycan; Signal; Alternative splicing;
KW   Pyrrolidone carboxylic acid.
FT   SIGNAL        1     20       BY SIMILARITY.
FT   CHAIN        21    362       CD44 ANTIGEN.
FT   DOMAIN       21    269       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    270    290       POTENTIAL.
FT   DOMAIN       291   362       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN        48    119      LINK.
FT   DOMAIN       150   158       ARG/LYS-RICH (BASIC).
FT   DOMAIN       225   269       STEM.
FT   MOD_RES       21     21      PYRROLIDONE CARBOXYLIC ACID (PROBABLE).
FT   DISULFID      53    118      BY SIMILARITY.
FT   DISULFID      77     97      BY SIMILARITY.
FT   MOD_RES      292    292      PHOSPHORYLATION (BY PKC) (PARTIAL) (BY
FT                                SIMILARITY).
FT   MOD_RES      326    326      PHOSPHORYLATION (PARTIAL) (BY
FT                                SIMILARITY).
FT   CARBOHYD      25     25      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      57     57      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     100    100      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     110    110      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     120    120      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     256    256      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT      67     67      E -> V (IN REF. 1; AA SEQUENCE).
SQ   SEQUENCE   362 AA;  39378 MW;  578BFCE7C3D52FFF CRC64;

  Query Match             13.4%;  Score 222.5;  DB 1;  Length 362;
  Best Local Similarity   25.5%;  Pred. No. 7.7e-11;
  Matches   92;  Conservative   43;  Mismatches  127;  Indels   99;  Gaps   16;

Qy          20 LVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQVETALK 79
               ||| ||   :: : ::||  ||  | |     :: ||| : |:    :|    |:| ||
Db          14 LVQLSL--AQIDLNITCRFEGIYHVEKNGRYSISRTEAADLCKAFNSTLPTMAQMEKALS 71

Qy          80 ASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDTWTNSC 139
                |||| ||:: :| ||| || ||  |  | || |     | |:  ||:|:|      |
Db          72 IGFETCRYGFI-EGHVVIPRIHPNSICAANNTGVYILTSNTS-QYDTYCFNASAPPGEDC 129

Qy         140 IPEIITTKDPIFNTQTATQTTEFIVSDST-YSVASPYSTIP------APT---------- 182
                    |:   : |      |   :  | | |     | | |      :||
```

```
Db        130 -----TSVTDLPNAFDGPITITIVNRDGTRYVKKGEYRTNPEDINPSSPTDDDVSSGSSS 184

Qy        183 ----------------TTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAF 226
                              |:|| |                 | ||:  |   :: ||
Db        185 ERSSTLGGYIFYNHFSTSPPIPDEDG--------------PWITDSTDRTPATRDQGAF 229

Qy        227 K----------NEAAGF------GG------------VPTALLVLALLFFGAAAGLGFCY 258
                      :|:||     ||            :|  |::|| |    |  |  |
Db        230 DPSGGSHTTHGSESAGHSHGSREGGANTTSGPLRTPQIPEWLIILASL-LALALILAVC- 287

Qy        259 VKRYVKAFPFTNKNQQKEMIETK---VVKEEKANDSNPNEESKKTDK----NPEESKSPS 311
                  :          |:|:::       |:: |::  |  | ||  :     | | |::|
Db        288 ----IAVNSRRRCGQKKKLVINNGNGAVEDRKSSGLN-GEASKSQEMVHLVNKESSETPD 342

Qy        312 K 312
                :
Db        343 Q 343


RESULT 3
CD44_RAT
ID   CD44_RAT       STANDARD;      PRT;   503 AA.
AC   P26051; Q99021;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   CD44 antigen precursor (Phagocytic glycoprotein I) (PGP-1) (HUTCH-I)
DE   (Extracellular matrix receptor-III) (ECMR-III) (GP90 lymphocyte
DE   homing/adhesion receptor) (Hermes antigen) (Hyaluronate receptor)
DE   (LY-24).
GN   CD44.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC   STRAIN=BDIX; TISSUE=Pancreas;
RX   MEDLINE=91191552; PubMed=1707342;
RA   Guenthert U., Hofmann M., Rudy W., Reber S., Zoeller M.,
RA   Haussmann I., Matzku S., Wenzel A., Ponta H., Herrlich P.;
RT   "A new variant of glycoprotein CD44 confers metastatic potential to
RT   rat carcinoma cells.";
RL   Cell 65:13-24(1991).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RA   Stevens J.W., Midura R.J.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Main cell surface receptor for hyaluronate. Adhesion to
CC       mucosal high endothelial venule and to types I and VI collagen.
CC       Probably involved in matrix adhesion, lymphocyte activation and
CC       lymph node homing.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC         Comment=Additional isoforms seem to exist;
CC       Name=2; Synonyms=Long, Meta-1;
CC         IsoId=P26051-1; Sequence=Displayed;
CC       Name=1; Synonyms=Short;
CC         IsoId=P26051-2; Sequence=VSP_005330;
CC   -!- PTM: N-glycosylated (By similarity).
CC   -!- PTM: O-glycosylated; contains chondroitin sulfate glycans which
CC       can be more or less sulfated (By similarity).
CC   -!- PTM: Phosphorylated; activation of PKC results in the
CC       dephosphorylation of Ser-467 (constitutive phosphorylation site),
CC       and the phosphorylation of Ser-433 (By similarity).
CC   -!- SIMILARITY: Contains 1 link domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M61875; AAA53532.1; -.
DR   EMBL; M61874; AAA53534.1; -.
DR   EMBL; U52179; AAA97915.1; -.
DR   EMBL; U46957; AAA92920.1; -.
DR   PIR; B38745; B38745.
DR   HSSP; P98066; 1TSG.
DR   InterPro; IPR001231; CD44_antigen.
DR   InterPro; IPR000538; Link.
DR   Pfam; PF00193; Xlink; 1.
DR   PRINTS; PR00658; CD44.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 1.
DR   SMART; SM00445; LINK; 1.
DR   PROSITE; PS01241; LINK; 1.
KW   Cell adhesion; Transmembrane; Glycoprotein; Phosphorylation; Receptor;
KW   Proteoglycan; Signal; Alternative splicing;
KW   Pyrrolidone carboxylic acid.
FT   SIGNAL       1     21
FT   CHAIN       22    503       CD44 ANTIGEN.
FT   DOMAIN      22    410       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   411    431       POTENTIAL.
FT   DOMAIN     432    503       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      51    123       LINK.
FT   DOMAIN     154    162       ARG/LYS-RICH (BASIC).
FT   DOMAIN     228    410       STEM.
FT   MOD_RES     22     22       PYRROLIDONE CARBOXYLIC ACID (PROBABLE).
FT   DISULFID    56    122       BY SIMILARITY.
FT   DISULFID    80    100       BY SIMILARITY.
FT   MOD_RES    433    433       PHOSPHORYLATION (BY PKC) (PARTIAL) (BY
FT                               SIMILARITY).
FT   MOD_RES    467    467       PHOSPHORYLATION (PARTIAL) (BY
FT                               SIMILARITY).
FT   CARBOHYD    28     28       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    60     60       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   103    103       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   114    114       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   124    124       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   266    266       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   274    274       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   306    306       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC   224    385       IATTPWVSAHTKQNQERTQWNPIHSNPEVLLQTTTRMTDID
FT                               RNSTSAHGENWTQEPQPPFNNHEYQDEEETPHATSTTWADP
FT                               NSTTEEAATQKEKWFENEWQGKNPPTPSEDSHVTEGTTASA
FT                               HNNHPSQRMTTQSQEDVSWTDFFDPISHPMGQGHQTESK
FT                               -> SDGDSSMDPRGGFDTVTHGSELA (in isoform
FT                               1).
FT                               /FTId=VSP_005330.
FT   CONFLICT    74     74       R -> S (IN REF. 2).
SQ   SEQUENCE   503 AA;  55945 MW;  FB489D009BD4EE22 CRC64;

  Query Match             13.4%;  Score 222;  DB 1;  Length 503;
  Best Local Similarity   25.7%;  Pred. No. 1.3e-10;
  Matches   84;  Conservative   40;  Mismatches  165;  Indels   38;  Gaps   11;

Qy         15 WTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQV 74
                |   |:| ||  ::: : ::||  |:  | |     :: ||| : |    :|    |:
Db         10 WGLLCLLQLSLAQQQIDLNITCRYAGVFHVEKNGRYSISRTEAADLCEAFNTTLPTMAQM 69

Qy         75 ETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDT 134
                | ||:  |||| ||:: :| ||| || ||  |  |  || |      :  :  ||:|:|
Db         70 ELALRKGFETCRYGFI-EGHVVIPRIHPNAICAANNTGVYILLASNTSHYDTYCFNASAP 128

Qy        135 WTNSCIPEIITTKDPIFNTQTATQTTEFIVSDST-YSVASPYST----IPAPT------- 182
                     |     |:   : |:     |  :  | | ||    | |    | |
Db        129 LEEDC-----TSVTDLPNSFDGPVTITIVNRDGTRYSKKGEYRTHQEDIDASNIIDEDVS 183

Qy        183 -------TTPPA-PASTSIPRRKKLICVTEVFMETSTMST-ETEPFVENKAAFKNEAAGF 233
                    :||   .  | :|  :      : |   ||::|  | |:|         |   :
Db        184 SGSTIEKSTPEGYILHTDLPTSQPTGDRDDAFFIGSTLATIATTPWVSAHTKQNQERTQW 243
```

```
Qy        234 GGV---PTALLVLALLFFGAAAGLGFCYVKRYVK--AFPFTNKNQQKEMIETKVVKEEKA 288
                 :    |  ||                 : : : :    || |   | |   ||
Db        244 NPIHSNPEVLLQTTTRMTDIDRNSTSAHGENWTQEPQPPFNNHEYQDEE-ETPHATSTTW 302

Qy        289 NDSNPNEESKKTDK-----NPEESKSP 310
                 | |    |   | |       |  : |:|
Db        303 ADPNSTTEEAATQKEKWFENEWQGKNP 329

RESULT 4
CD44_MESAU
ID   CD44_MESAU     STANDARD;      PRT;   431 AA.
AC   Q60522; Q60523;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   CD44 antigen precursor (Phagocytic glycoprotein I) (PGP-1) (HUTCH-I)
DE   (Extracellular matrix receptor-III) (ECMR-III) (GP90 lymphocyte
DE   homing/adhesion receptor) (Hermes antigen) (Hyaluronate receptor)
DE   (Heparan sulfate proteoglycan) (HAM1 antigen).
GN   CD44.
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Mesocricetus.
OX   NCBI_TaxID=10036;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC   STRAIN=LVG; TISSUE=Alveolar macrophage;
RA   Paulauskis J.D., Kobzik L., Gerard C., Katler M., Godleski J.J.;
RL   Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Main cell surface receptor for hyaluronate. Adhesion to
CC       mucosal high endothelial venule and to types I and VI collagen.
CC       Probably involved in matrix adhesion, lymphocyte activation and
CC       lymph node homing.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC         Comment=Additional isoforms seem to exist;
CC       Name=1;
CC         IsoId=Q60522-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q60522-2; Sequence=VSP_005322;
CC   -!- PTM: N-glycosylated (By similarity).
CC   -!- PTM: O-glycosylated; contains chondroitin sulfate glycans which
CC       can be more or less sulfated (By similarity).
CC   -!- PTM: Phosphorylated; activation of PKC results in the
CC       dephosphorylation of Ser-395 (constitutive phosphorylation site),
CC       and the phosphorylation of Ser-361 (By similarity).
CC   -!- SIMILARITY: Contains 1 link domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U10880; AAC13767.1; -.
DR   EMBL; U10881; AAA19316.1; -.
DR   HSSP; P98066; 1TSG.
DR   InterPro; IPR001231; CD44_antigen.
DR   InterPro; IPR000538; Link.
DR   Pfam; PF00193; Xlink; 1.
DR   PRINTS; PR00658; CD44.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 1.
DR   SMART; SM00445; LINK; 1.
DR   PROSITE; PS01241; LINK; 1.
KW   Cell adhesion; Transmembrane; Glycoprotein; Phosphorylation; Receptor;
KW   Proteoglycan; Signal; Alternative splicing;
KW   Pyrrolidone carboxylic acid.
FT   SIGNAL        1     22       BY SIMILARITY.
FT   CHAIN        23    431       CD44 ANTIGEN.
FT   DOMAIN       23    338       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    339    359       POTENTIAL.
FT   DOMAIN      360    431       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       50    121       LINK.
FT   DOMAIN      152    160       ARG/LYS-RICH (BASIC).
FT   DOMAIN      226    338       STEM.
FT   MOD_RES      23     23       PYRROLIDONE CARBOXYLIC ACID (BY
FT                                SIMILARITY).
FT   DISULFID     55    120       BY SIMILARITY.
FT   DISULFID     79     99       BY SIMILARITY.
FT   MOD_RES     361    361       PHOSPHORYLATION (BY PKC) (PARTIAL) (BY
FT                                SIMILARITY).
FT   MOD_RES     395    395       PHOSPHORYLATION (PARTIAL) (BY
FT                                SIMILARITY).
FT   CARBOHYD     27     27       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     59     59       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    102    102       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    112    112       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    122    122       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    174    174       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    256    256       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    325    325       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    222    291       TRSGGKDGRRGGGLPKDATTSLEGYTTHYPETMENGTLTPV
FT                                TPAKTGVFGETEVTVAEDSNFNVDGSLPG -> R (in
FT                                isoform 2).
FT                                /FTId=VSP_005322.
SQ   SEQUENCE   431 AA;  46807 MW;  4300262E0C6BEA6A CRC64;

  Query Match             12.8%;  Score 212.5;  DB 1;  Length 431;
  Best Local Similarity   25.7%;  Pred. No. 6.3e-10;
  Matches   80;  Conservative   40;  Mismatches  128;  Indels   63;  Gaps   12;

Qy         24 SLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQVETALKASFE 83
              ||  ::: : ::||  |:  |  |      :: |||  : |:      :|   ||:  ||   ||
Db         18 SLAQQQIDLNITCRYAGVFHVEKNGRYSISRTEAADLCQAFNSTLPTMDQMVMALSKGFE 77

Qy         84 TCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDTWTNSCIPEI 143
              || ||:: :| |||  || ||   |   |   ||     :   ||:|:|      |
Db         78 TCRYGFI-EGHVVIPRIQPNAICAANHTGVYILTSNTS-HYDTYCFNASAPLEEDC---- 131

Qy        144 ITTKDPIFNTQTATQTTEFIVSDST-YSVASPYST----IPAPTTTPPAPAS-------- 190
              |:    : |:      |    :  | | ||    |  |     | |  ||      :|
Db        132 -TSVTDLPNSFEGPVTITIVNRDGTRYSKKGEYRTHQEDIDASNTTDDDVSSGSSSEKST 190

Qy        191 -------TSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTALLVL 243
                     | :|          : :   |||:|          :: |:   | ||:|
Db        191 SGGYVFHTYLPTIHSTADQDDPYFIGSTMAT-------TRSGGKDGRRG-GGLPK----- 237

Qy        244 ALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMI---------ETKVVKEEKAN----D 290
                   |     |        : |    :| | :|       :           ||:|    | :|
Db        238 -----DATTSL-----EGYTTHYPETMENGTLTPVTPAKTGVFGETEVTVAEDSNFNVDG 287

Qy        291 SNPNEESKKTD 301
              | | ::   |
Db        288 SLPGDQDSSMD 298

RESULT 5
CD44_MOUSE
ID   CD44_MOUSE     STANDARD;      PRT;   778 AA.
AC   P15379; Q05732; Q61395; Q62060; Q62061; Q62062; Q62063; Q62408;
AC   Q62409; Q64296; Q99J14; Q9QYX8;
DT   01-APR-1990 (Rel. 14, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   CD44 antigen precursor (Phagocytic glycoprotein I) (PGP-1) (HUTCH-I)
DE   (Extracellular matrix receptor-III) (ECMR-III) (GP90 lymphocyte
DE   homing/adhesion receptor) (Hermes antigen) (Hyaluronate receptor)
DE   (LY-24).
GN   CD44.
OS   Mus musculus (Mouse).
```

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 4; 6; 7 AND 12).
RC   STRAIN=DBA/2; TISSUE=Lung;
RX   MEDLINE=93107170; PubMed=1469058;
RA   He Q., Lesley J., Hyman R., Ishihara K., Kincade P.W.;
RT   "Molecular isoforms of murine CD44 and evidence that the membrane
RT   proximal domain is not critical for hyaluronate recognition.";
RL   J. Cell Biol. 119:1711-1719(1992).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 13).
RX   MEDLINE=90038499; PubMed=2681416;
RA   Zhou D.F.H., Ding J.F., Picker L.J., Bargatze R.F., Butcher E.C.,
RA   Goeddel D.V.;
RT   "Molecular cloning and expression of Pgp-1. The mouse homolog of the
RT   human H-CAM (Hermes) lymphocyte homing receptor.";
RL   J. Immunol. 143:3390-3395(1989).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 13).
RX   MEDLINE=90046829; PubMed=2682651;
RA   Nottenburg C., Rees G., St John T.;
RT   "Isolation of mouse CD44 cDNA: structural features are distinct from
RT   the primate cDNA.";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:8521-8525(1989).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RX   MEDLINE=20318634; PubMed=10859330;
RA   Wittig B.M., Johansson B., Zoeller M., Schwaerzler C., Guenthert U.;
RT   "Abrogation of experimental colitis correlates with increased
RT   apoptosis in mice deficient for CD44 variant exon 7 (CD44v7).";
RL   J. Exp. Med. 191:2053-2064(2000).
RN   [5]
RP   SEQUENCE FROM N.A. (ISOFORM 13).
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [6]
RP   SEQUENCE FROM N.A. (ISOFORM 13).
RC   STRAIN=C57BL/6J; TISSUE=Embryo;
RX   MEDLINE=22354683; PubMed=12466851;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA   Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA   Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA   Birney E., Hayashizaki Y.;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [7]
RP   SEQUENCE OF 8-778 FROM N.A. (ISOFORM 13).
RX   MEDLINE=90094420; PubMed=2403559;
RA   Wolffe E.J., Gause W.C., Pelfrey C.M., Holland S.M., Steinberg A.D.,
RA   August J.T.;
RT   "The cDNA sequence of mouse Pgp-1 and homology to human CD44 cell
RT   surface antigen and proteoglycan core/link proteins.";
RL   J. Biol. Chem. 265:341-347(1990).
RN   [8]
RP   SEQUENCE OF 224-637 FROM N.A. (ISOFORMS 1; 2; 3; 4; 5; 6; 7 AND 8).
RC   STRAIN=GR;
RX   MEDLINE=93219085; PubMed=8464707;
RA   Toelg C., Hofmann M., Herrlich P., Ponta H.;
RT   "Splicing choice from ten variant exons establishes CD44
RT   variability.";
RL   Nucleic Acids Res. 21:1225-1229(1993).
RN   [9]
RP   SEQUENCE OF 224-637 FROM N.A. (ISOFORM 9).
RC   STRAIN=BALB/c;
RX   MEDLINE=93286043; PubMed=8509359;
RA   Screaton G.R., Bell M.V., Bell J.I., Jackson D.G.;
RT   "The identification of a new alternative exon with highly restricted
RT   tissue expression in transcripts encoding the mouse Pgp-1 (CD44)
RT   homing receptor. Comparison of all 10 variable exons between mouse,
RT   human, and rat.";
RL   J. Biol. Chem. 268:12235-12238(1993).
RN   [10]
RP   PARTIAL SEQUENCE FROM N.A. (ISOFORMS 10 AND 11).
RC   STRAIN=Swiss Webster;
RX   MEDLINE=96355396; PubMed=8702806;
RA   Yu Q., Toole B.P.;
RT   "A new alternatively spliced exon between v9 and v10 provides a
RT   molecular basis for synthesis of soluble CD44.";
RL   J. Biol. Chem. 271:20603-20607(1996).
CC   -!- FUNCTION: Main cell surface receptor for hyaluronate. Adhesion to
CC       mucosal high endothelial venule and to types I and VI collagen.
CC       Probably involved in matrix adhesion, lymphocyte activation and
CC       lymph node homing.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=13;
CC       Name=1;
CC         IsoId=P15379-14; Sequence=Displayed;
CC       Name=2;
CC         IsoId=P15379-7; Sequence=VSP_007329;
CC       Name=3;
CC         IsoId=P15379-8; Sequence=VSP_007330;
CC       Name=4; Synonyms=M2;
CC         IsoId=P15379-4; Sequence=VSP_007331;
CC       Name=5;
CC         IsoId=P15379-9; Sequence=VSP_007332;
CC       Name=6; Synonyms=M3;
CC         IsoId=P15379-5; Sequence=VSP_005326;
CC       Name=7; Synonyms=M4;
CC         IsoId=P15379-6; Sequence=VSP_005327;
CC       Name=8;
CC         IsoId=P15379-10; Sequence=VSP_007330, VSP_007334;
```

```
CC       Name=9;
CC         IsoId=P15379-11; Sequence=VSP_007332, VSP_007335;
CC       Name=10;
CC         IsoId=P15379-12; Sequence=VSP_007336, VSP_007337;
CC       Name=11;
CC         IsoId=P15379-13; Sequence=VSP_007338, VSP_007339;
CC       Name=12; Synonyms=M1;
CC         IsoId=P15379-3; Sequence=VSP_005328;
CC       Name=13; Synonyms=M0;
CC         IsoId=P15379-2; Sequence=VSP_005329;
CC   -!- PTM: N-glycosylated (By similarity).
CC   -!- PTM: O-glycosylated; contains chondroitin sulfate glycans which
CC       can be more or less sulfated (By similarity).
CC   -!- PTM: Phosphorylated; activation of PKC results in the
CC       dephosphorylation of Ser-742 (constitutive phosphorylation site),
CC       and the phosphorylation of Ser-708 (By similarity).
CC   -!- POLYMORPHISM: Two allelic forms of this glycoprotein, PGP-1.1 and
CC       PGP-1.2, have been reported. The expressed product is PGP-1.1 (Ly-
CC       24.1).
CC   -!- SIMILARITY: Contains 1 link domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X66084; CAA46883.1; -.
DR   EMBL; X66083; CAA46882.1; -.
DR   EMBL; X66082; CAA46881.1; -.
DR   EMBL; X66081; CAA46880.1; -.
DR   EMBL; M30655; AAA39922.1; -.
DR   EMBL; M27129; AAA37406.1; -.
DR   EMBL; M27130; AAA37407.1; -.
DR   EMBL; AJ251594; CAB61888.1; -.
DR   EMBL; BC005676; AAH05676.1; -.
DR   EMBL; AK045226; BAC32269.1; -.
DR   EMBL; J05163; AAA39923.1; -.
DR   EMBL; X69724; CAA49380.1; -.
DR   EMBL; L13611; AAA37145.1; -.

  Query Match             12.7%;  Score 210;  DB 1;  Length 778;
  Best Local Similarity   32.3%;  Pred. No. 2.1e-09;
  Matches   54;  Conservative   25;  Mismatches   78;  Indels   10;  Gaps    4;

Qy         15 WTTR---LLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGK 71
                 | |      |:| ||  ::: : |:||  |:  | |     :: ||| : |:    :|
Db          6 WHTAWGLCLLQLSLAHQQIDLNVTCRYAGVFHVEKNGRYSISRTEAADLCQAFNSTLPTM 65

Qy         72 DQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNS 131
              ||:: ||   |||| ||:: :| ||| || ||  |  |  || |     :  :  ||:|:
Db         66 DQMKLALSKGFETCRYGFI-EGNVVIPRIHPNAICAANHTGVYILVTSNTSHYDTYCFNA 124

Qy        132 SDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDST-YSVASPYST 177
                  |       |     |:   : |:     |   :  | | ||    | |
Db        125 SAPPEEDC-----TSVTDLPNSFDGPVTITIVNRDGTRYSKKGEYRT 166


RESULT 6
CD44_BOVIN
ID   CD44_BOVIN     STANDARD;      PRT;   366 AA.
AC   Q29423;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   CD44 antigen precursor (Phagocytic glycoprotein I) (PGP-1) (HUTCH-I)
DE   (Extracellular matrix receptor-III) (ECMR-III) (GP90 lymphocyte
DE   homing/adhesion receptor) (Hermes antigen) (Hyaluronate receptor).
GN   CD44.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92017904; PubMed=1922105;
RA   Bosworth B.T., St John T., Gallatin W.M., Harp J.A.;
RT   "Sequence of the bovine CD44 cDNA: comparison with human and mouse
RT   sequences.";
RL   Mol. Immunol. 28:1131-1135(1991).
CC   -!- FUNCTION: Main cell surface receptor for hyaluronate. Adhesion to
CC       mucosal high endothelial venule and to types I and VI collagen.
CC       Probably involved in matrix adhesion, lymphocyte activation and
CC       lymph node homing.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: Mesenteric lymph node and liver, not in heart.
CC   -!- PTM: Extensively modified including N- and O-linked glycosylation,
CC       addition of the glycosaminoglycan chondroitin sulfate, of sulfate,
CC       of phosphate to cytoplasmic domain serine residues (By
CC       similarity).
CC   -!- SIMILARITY: Contains 1 link domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X62881; CAA44675.1; -.
DR   EMBL; S63418; AAB20016.1; -.
DR   PIR; A53286; A53286.
DR   HSSP; P98066; 1TSG.
DR   InterPro; IPR001231; CD44_antigen.
DR   InterPro; IPR000538; Link.
DR   Pfam; PF00193; Xlink; 1.
DR   PRINTS; PR00658; CD44.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 1.
DR   SMART; SM00445; LINK; 1.
DR   PROSITE; PS01241; LINK; 1.
KW   Cell adhesion; Transmembrane; Glycoprotein; Phosphorylation; Receptor;
KW   Proteoglycan; Signal; Alternative splicing;
KW   Pyrrolidone carboxylic acid.
FT   SIGNAL        1     20       POTENTIAL.
FT   CHAIN        21    366       CD44 ANTIGEN.
FT   DOMAIN       21    273       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    274    294       POTENTIAL.
FT   DOMAIN      295    366       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       48    119       LINK.
FT   DOMAIN      150    158       ARG/LYS-RICH (BASIC).
FT   DOMAIN      229    273       STEM.
FT   MOD_RES      21     21       PYRROLIDONE CARBOXYLIC ACID (BY
FT                                SIMILARITY).
FT   DISULFID     53    118       BY SIMILARITY.
FT   DISULFID     77     97       BY SIMILARITY.
FT   MOD_RES     296    296       PHOSPHORYLATION (BY PKC) (PARTIAL) (BY
FT                                SIMILARITY).
FT   MOD_RES     330    330       PHOSPHORYLATION (PARTIAL) (BY
FT                                SIMILARITY).
FT   CARBOHYD     25     25       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     57     57       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    100    100       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    120    120       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    222    222       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    260    260       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   366 AA;  40001 MW;  438A5A1E631E02B4 CRC64;

  Query Match             12.6%;  Score 208.5;  DB 1;  Length 366;
  Best Local Similarity   23.8%;  Pred. No. 1.1e-09;
  Matches   81;  Conservative   44;  Mismatches  142;  Indels   73;  Gaps   11;
```

```
Qy          20 LVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQVETALK 79
               ||| ||   :: : ::||  |:  | |     :: ||| : |:    :|    |:| |
Db          14 LVQLSL--AQIDLNITCRYAGVFHVEKNGRYSISKTEAADLCKAFNSTLPTMAQMEAARN 71

Qy          80 ASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDTWTNSC 139
                |||| ||:: :| ||| || ||  |  |  || |     | |:   |:|:|      |
Db          72 IGFETCRYGFI-EGHVVIPRIHPNSICAANNTGVYILTSNTS-QYDTICFNASAPPGEDC 129

Qy         140 IPEIITTKDPIFNTQTATQTTEFIVSDST-YSVASPYSTIP--------APTTTPPAPAS 190
                 |:   : |      |   :  | | |:    | | |         :|:: |    |
Db         130 -----TSVTDLPNAFEGPITITIVNRDGTRYTKKGEYRTNPEDINPSVVSPSSPPDDEMS 184

Qy         191 TSIPRRKKLICVTEVF----------------------METSTMSTETEPFVENKAAFK 227
               :  |  :       :|                     :|    :  :|   :
Db         185 SGSPSERSTSGGYSIFHTHLPTVHPSRPRRPWSQRAENTSDTRDYGSSHDPSGRSYTTHA 244

Qy         228 NEAAGFGG-----------------VPTALLVLALLFFGAAAGLGFCYVKRYVKAFPFT 269
               :|:||                  :|  |::|| |    |  | |    :
Db         245 SESAGHSSGSEEHGANTTSGPMRKPQIPEWLIILASL-LALALILAVC-----IAVNSRR 298

Qy         270 NKNQQKEMIETKVVKEEKANDSNPNEESKKTDKNPEESKS 309
                  |:|:::          | :   || | :  | | |||
Db         299 RCGQKKKLVIN--------NGNGTMEERKPSGLNGEASKS 330
```

```
RESULT 7
CD44_CANFA
ID   CD44_CANFA     STANDARD;      PRT;   351 AA.
AC   Q28284;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   CD44 antigen precursor (Phagocytic glycoprotein I) (PGP-1) (HUTCH-I)
DE   (Extracellular matrix receptor-III) (ECMR-III) (GP90 lymphocyte
DE   homing/adhesion receptor) (Hermes antigen) (Hyaluronate receptor)
DE   (Fragment).
GN   CD44.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Beagle; TISSUE=Thymus;
RX   MEDLINE=94250687; PubMed=7514890;
RA   Milde K.F., Alejandro R., Mintz D.H., Pastori R.L.;
RT   "Molecular cloning of the canine CD44 antigen cDNA.";
RL   Biochim. Biophys. Acta 1218:112-114(1994).
CC   -!- FUNCTION: Main cell surface receptor for hyaluronate. Adhesion to
CC       mucosal high endothelial venule and to types I and VI collagen.
CC       Probably involved in matrix adhesion, lymphocyte activation and
CC       lymph node homing.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: Lymph nodes.
CC   -!- PTM: Extensively modified including N- and O-linked glycosylation,
CC       addition of the glycosaminoglycan chondroitin sulfate, of sulfate,
CC       of phosphate to cytoplasmic domain serine residues (By
CC       similarity).
CC   -!- SIMILARITY: Contains 1 link domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z27115; CAA81630.1; -.
DR   PIR; S45305; S45305.
DR   HSSP; P98066; 1TSG.
DR   InterPro; IPR001231; CD44_antigen.
DR   InterPro; IPR000538; Link.
DR   Pfam; PF00193; Xlink; 1.
DR   PRINTS; PR00658; CD44.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 1.
DR   SMART; SM00445; LINK; 1.
DR   PROSITE; PS01241; LINK; 1.
KW   Cell adhesion; Transmembrane; Glycoprotein; Phosphorylation; Receptor;
KW   Proteoglycan; Signal; Alternative splicing;
KW   Pyrrolidone carboxylic acid.
FT   NON_TER       1      1
FT   SIGNAL       <1     13       POTENTIAL.
FT   CHAIN         14   >351       CD44 ANTIGEN.
FT   DOMAIN        14    263       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     264    284       POTENTIAL.
FT   DOMAIN       285   >351       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN        41    112       LINK.
FT   DOMAIN       143    151       ARG/LYS-RICH (BASIC).
FT   DOMAIN       218    263       STEM.
FT   MOD_RES       14     14       PYRROLIDONE CARBOXYLIC ACID (BY
FT                                 SIMILARITY).
FT   DISULFID      46    111       BY SIMILARITY.
FT   DISULFID      70     90       BY SIMILARITY.
FT   MOD_RES      286    286       PHOSPHORYLATION (BY PKC) (PARTIAL) (BY
FT                                 SIMILARITY).
FT   MOD_RES      320    320       PHOSPHORYLATION (PARTIAL) (BY
FT                                 SIMILARITY).
FT   CARBOHYD      18     18       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      50     50       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      93     93       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     103    103       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     113    113       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     250    250       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   NON_TER      351    351
SQ   SEQUENCE   351 AA;  38066 MW;  E73387E70E20C0E0 CRC64;
```

```
  Query Match             12.2%;  Score 202.5;  DB 1;  Length 351;
  Best Local Similarity   23.4%;  Pred. No. 3.2e-09;
  Matches   84;  Conservative   49;  Mismatches  145;  Indels   81;  Gaps   15;

Qy          15 WTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQV 74
               |    ||    |   :: : ::||  |:  | |     :: | | : |:    :|    |:
Db           3 WGLCLL---RLSLAQIDLNITCRYAGVFHVEKNGRYSISRTAAADLCKAFNSTLPTMAQM 59

Qy          75 ETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDT 134
               | ||   |||| ||:: :| ||| || ||  |  | || | :  : |:  ||:|:|
Db          60 ERALSVGFETCRYGFI-EGHVVIPRIQPNAICAANHTGVYI-LISNTSQYDTYCFNASAP 117

Qy         135 WTNSCIPEIITTKDPIFNTQTATQTTEFIVSDST-YSVASPYSTIPA---PTTTPPAPAS 190
                    |    |:   : |       |    :  | | ||    | | |    |:       |
Db         118 PEEDC-----TSVTHLPNAFDGPITITIVNRDGTRYSQKGEYRTNPEDINPSNPTDDDVS 172

Qy         191 TSIPRRKKLICVTEVF------------METSTMSTETE-------------PFVENKAA 225
               :     :       :|              ::| :|: ::             |   :
Db         173 SGSSSERSTSAGYNIFHTHLPTAYPTEDQDSSRVSSNSDHTPITKDHDSSVHPSERSHTT 232

Qy         226 FKNEAAGF------GG------------VPTALLVLALLFFGAAAGLGFCYVKRYVKAFP 267
               :|:||       ||            :|  |::|| |    | | |
Db         233 HGSESAGHSSGSQEGGANTTSGPMRKPQIPEWLIILASL-LALALILAVCIA-------- 283

Qy         268 FTNKNQQKEMIETKVVKEEKAN----DSNP---NEESKKTDK-----NPEESKSPSKTT 314
                  | ::   : | :      |    |  |   | |: |: :     | | |::| : |
Db         284 ---VNSRRRCGQKKKLVINNGNGAVGDRKPSGINGEASKSQEMVHLVNKEPSETPDQYT 339
```

```
RESULT 8
CD44_HUMAN
ID   CD44_HUMAN     STANDARD;      PRT;   742 AA.
AC   P16070; P22511; Q04858; Q13419; Q13957; Q13958; Q13959; Q13960;
AC   Q13961; Q13967; Q13968; Q13980; Q15861; Q16064; Q16065; Q16066;
AC   Q16208; Q16522; Q96J24;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
```

```
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   CD44 antigen precursor (Phagocytic glycoprotein I) (PGP-1) (HUTCH-I)
DE   (Extracellular matrix receptor-III) (ECMR-III) (GP90 lymphocyte
DE   homing/adhesion receptor) (Hermes antigen) (Hyaluronate receptor)
DE   (Heparan sulfate proteoglycan) (Epican) (CDw44).
GN   CD44 OR LHR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC   TISSUE=Lymphoblast;
RX   MEDLINE=93101687; PubMed=1465456;
RA   Screaton G.R., Bell M.V., Jackson D.G., Cornelis F.B., Gerth U.,
RA   Bell J.I.;
RT   "Genomic structure of DNA encoding the lymphocyte homing receptor
RT   CD44 reveals at least 12 alternatively spliced exons.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:12160-12164(1992).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM EPIDERMAL).
RC   TISSUE=Keratinocytes;
RA   Kugelman L.C., Ganguly S., Haggerty J.G., Weissman S.M.,
RA   Milstone L.M.;
RT   "The core protein of epican, a heparan sulfate proteoglycan on
RT   keratinocytes, is an alternative form of CD44.";
RL   J. Invest. Dermatol. 99:381-385(1992).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM EPITHELIAL CD44E).
RX   MEDLINE=91122041; PubMed=1991450;
RA   Stamenkovic I., Aruffo A., Amiot M., Seed B.;
RT   "The hematopoietic and epithelial forms of CD44 are distinct
RT   polypeptides with different adhesion potentials for
RT   hyaluronate-bearing cells.";
RL   EMBO J. 10:343-348(1991).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORMS CD44R1 AND CD44R2).
RC   TISSUE=Myeloid leukemia cells;
RX   MEDLINE=91277598; PubMed=2056274;
RA   Dougherty G.J., Lansdorp P.M., Cooper D.L., Humphries R.K.;
RT   "Molecular cloning of CD44R1 and CD44R2, two novel isoforms of the
RT   human CD44 lymphocyte 'homing' receptor expressed by hemopoietic
RT   cells.";
RL   J. Exp. Med. 174:1-5(1991).
RN   [5]
RP   SEQUENCE FROM N.A. (ISOFORMS CD44SP; CD44R4 AND CD44R5).
RC   TISSUE=Breast carcinoma;
RX   MEDLINE=93356912; PubMed=8352881;
RA   Tanabe K.K., Nishi T., Saya H.;
RT   "Novel variants of CD44 arising from alternative splicing: changes in
RT   the CD44 alternative splicing pattern of MCF-7 breast carcinoma cells
RT   treated with hyaluronidase.";
RL   Mol. Carcinog. 7:212-220(1993).
RN   [6]
RP   SEQUENCE FROM N.A. (ISOFORM RETICULOCYTE).
RC   TISSUE=Reticulocytes;
RX   MEDLINE=91337049; PubMed=1840487;
RA   Harn H.J., Isola N., Cooper D.L.;
RT   "The multispecific cell adhesion molecule CD44 is represented in
RT   reticulocyte cDNA.";
RL   Biochem. Biophys. Res. Commun. 178:1127-1134(1991).
RN   [7]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89168434; PubMed=2466575;
RA   Stamenkovic I., Amiot M., Pesando J.M., Seed B.;
RT   "A lymphocyte molecule implicated in lymph node homing is a member of
RT   the cartilage link protein family.";
RL   Cell 56:1057-1062(1989).
RN   [8]
RP   SEQUENCE FROM N.A. (ISOFORMS SHORT-TAILED HEMATOPOIETIC AND CD44H).
RX   MEDLINE=89168435; PubMed=2466576;
RA   Goldstein L.A., Zhou D.F.H., Picker L.J., Minty C.N., Bargatze R.F.,
RA   Ding J.F., Butcher E.C.;
RT   "A human lymphocyte homing receptor, the hermes antigen, is related
RT   to cartilage proteoglycan core and link proteins.";
RL   Cell 56:1063-1072(1989).
RN   [9]
RP   SEQUENCE FROM N.A. (ISOFORM WITHOUT EXON 6).
RC   TISSUE=Pancreas;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [10]
RP   SEQUENCE OF 184-222 AND 473-625 FROM N.A. (ISOFORM KERATINOCYTE).
RC   TISSUE=Foreskin;
RX   MEDLINE=91177958; PubMed=2007624;
RA   Brown T.A., Bouchard T., St John T., Wayner E., Carter W.G.;
RT   "Human keratinocytes express a new CD44 core protein (CD44E) as a
RT   heparan-sulfate intrinsic membrane proteoglycan with additional
RT   exons.";
RL   J. Cell Biol. 113:207-221(1991).
RN   [11]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Articular cartilage;
RA   Bosch P.P., Stevens J.W., Buckwalter J.A., Midura R.J.;
RL   Submitted (DEC-1995) to the EMBL/GenBank/DDBJ databases.
RN   [12]
RP   SEQUENCE OF 267-603 FROM N.A.
RC   TISSUE=Lung;
RX   MEDLINE=92005448; PubMed=1717145;
RA   Hofmann M., Rudy W., Zoeller M., Toelg C., Ponta H., Herrlich P.,
RA   Guenthert U.;
RT   "CD44 splice variants confer metastatic behavior in rats: homologous
RT   sequences are expressed in human tumor cell lines.";
RL   Cancer Res. 51:5292-5297(1991).
RN   [13]
RP   SEQUENCE OF 223-265 FROM N.A.
RX   MEDLINE=94198700; PubMed=8148709;
RA   Matsumura Y., Hanbury D., Smith J., Tarin D.;
RT   "Non-invasive detection of malignancy by identification of unusual
RT   CD44 gene activity in exfoliated cancer cells.";
RL   BMJ 308:619-624(1994).
RN   [14]
RP   SEQUENCE OF 1-22 FROM N.A.
RC   TISSUE=Lymphoblast;
RX   MEDLINE=92017823; PubMed=1922057;
RA   Shtivelman E., Bishop J.M.;
RT   "Expression of CD44 is repressed in neuroblastoma cells.";
RL   Mol. Cell. Biol. 11:5446-5453(1991).
RN   [15]
RP   REVIEW ON FUNCTION AND POST-TRANSLATIONAL MODIFICATIONS.
RX   MEDLINE=22399881; PubMed=12511867;
RA   Ponta H., Sherman L., Herrlich P.A.;
RT   "CD44: from adhesion molecules to signalling regulators.";
RL   Nat. Rev. Mol. Cell Biol. 4:33-45(2003).
RN   [16]
RP   PHOSPHORYLATION OF SER-706.
RX   MEDLINE=98248445; PubMed=9580567;
```

```
RA   Peck D., Isacke C.M.;
RT   "Hyaluronan-dependent cell migration can be blocked by a CD44
RT   cytoplasmic domain peptide containing a phosphoserine at position
RT   325.";
RL   J. Cell Sci. 111:1595-1601(1998).
RN   [17]
RP   PHOSPHORYLATION OF SER-672.
RX   MEDLINE=22038351; PubMed=12032545;
RA   Legg J.W., Lewis C.A., Parsons M., Ng T., Isacke C.M.;
RT   "A novel PKC-regulated mechanism controls CD44 ezrin association and
RT   directional cell motility.";
RL   Nat. Cell Biol. 4:399-407(2002).
RN   [18]
RP   GLYCOSYLATION AND PROCESSING.
RX   MEDLINE=22766015; PubMed=12883358;
RA   Bartolazzi A.;
RT   "CD44s adhesive function spontaneous and PMA-inducible CD44 cleavage
RT   are regulated at post-translational level in cells of melanocytic
RT   lineage.";
RL   Melanoma Res. 13:325-337(2003).
RN   [19]
RP   VARIANT BLOOD GROUP INDIAN PRO-46.
RX   MEDLINE=96215152; PubMed=8636151;
RA   Telen M.J., Udani M., Washington M.K., Levesque M.C., Lloyd E.,
RA   Rao N.;
RT   "A blood group-related polymorphism of CD44 abolishes a hyaluronan-
RT   binding consensus sequence without preventing hyaluronan binding.";
RL   J. Biol. Chem. 271:7147-7153(1996).
CC   -!- FUNCTION: Receptor for hyaluronic acid (HA). Mediates cell-cell
CC       and cell-matrix interactions through its affinity for HA, and
CC       possibly also through its affinity for other ligands such as
CC       osteopontin, collagens, and matrix matalloproteinases (MMPs).
CC       Adhesion with HA plays an important role in cell migration, tumor
CC       growth and progression. Also involved in lymphocyte activation,
CC       recirculation and homing, and in hematopoiesis. Altered expression
CC       or dysfunction causes numerous pathogenic phenotypes. Great
CC       protein heterogeneity due to numerous alternative splicing and
CC       post-translational modification events.
CC   -!- SUBUNIT: Interacts with HA, as well as other glycosaminoglycans,
CC       collagen, laminin, and fibronectin via its N-terminal segment.
CC       Interacts with ANK, the ERM proteins (VIL2, RDX and
CC       MSN), and NF2 via its C-terminal segment.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC         Comment=Additional isoforms seem to exist. Additional isoforms
CC         are produced by alternative splicing of 10 out of 19 exons
CC         within the extracellular domain. Additional diversity is
```

```
  Query Match             12.2%;  Score 202;  DB 1;  Length 742;
  Best Local Similarity   26.0%;  Pred. No. 8.8e-09;
  Matches   61;  Conservative   30;  Mismatches   84;  Indels   60;  Gaps    7;

Qy         25 LRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQVETALKASFET 84
              |   :: : ::||  |:  | |     :: ||| : |:    :|    |:| ||   |||
Db         17 LSLAQIDLNITCRFAGVFHVEKNGRYSISRTEAADLCKAFNSTLPTMAQMEKALSIGFET 76

Qy         85 CSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDTWTNSCI---- 140
              | ||:: :| ||| || ||  |  |  || |     | |:  ||:|:|      |
Db         77 CRYGFI-EGHVVIPRIHPNSICAANNTGVYILTSNTS-QYDTYCFNASAPPEEDCTSVTD 134

Qy        141 ------------------------------PEIITTKDPIFNTQTATQTTEFIVSDSTYS 170
                                           || |   :|  :  ::  ::|   :   |
Db        135 LPNAFDGPITITIVNRDGTRYVQKGEYRTNPEDIYPSNPTDDDVSSGSSSERSSTSGGY- 193

Qy        171 VASPYSTI-PAPTTTPP--------APASTSIPRRKKLICVTEVFMETSTMSTET 216
              :   :||: | |    |          ||:|               | ||  :|||
Db        194 IFYTFSTVHPIPDEDSPWITDSTDRIPATT-------------LMSTSATATET 234
```

```
RESULT 9
CD44_HORSE
ID   CD44_HORSE     STANDARD;      PRT;   359 AA.
AC   Q05078;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   CD44 antigen precursor (Phagocytic glycoprotein I) (PGP-1) (HUTCH-I)
DE   (Extracellular matrix receptor-III) (ECMR-III) (GP90 lymphocyte
DE   homing/adhesion receptor) (Hermes antigen) (Hyaluronate receptor).
GN   CD44.
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX   NCBI_TaxID=9796;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93170897; PubMed=8436424;
RA   Tavernor A.S., Deverson E.V., Coadwell W.J., Lunn D.P., Zhang C.,
RA   Davis W., Butcher G.W.;
RT   "Molecular cloning of equine CD44 cDNA by a COS cell expression
RT   system.";
RL   Immunogenetics 37:474-477(1993).
CC   -!- FUNCTION: Main cell surface receptor for hyaluronate. Adhesion to
CC       mucosal high endothelial venule and to types I and VI collagen.
CC       Probably involved in matrix adhesion, lymphocyte activation and
CC       lymph node homing.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- PTM: Extensively modified including N- and O-linked glycosylation,
CC       addition of the glycosaminoglycan chondroitin sulfate, of sulfate,
CC       of phosphate to cytoplasmic domain serine residues.
CC   -!- SIMILARITY: Contains 1 link domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X66862; CAA47331.1; -.
DR   PIR; I46245; S24240.
DR   HSSP; P98066; 1TSG.
DR   InterPro; IPR001231; CD44_antigen.
DR   InterPro; IPR000538; Link.
DR   Pfam; PF00193; Xlink; 1.
DR   PRINTS; PR00658; CD44.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 1.
DR   SMART; SM00445; LINK; 1.
DR   PROSITE; PS01241; LINK; 1.
KW   Cell adhesion; Transmembrane; Glycoprotein; Phosphorylation; Receptor;
KW   Proteoglycan; Signal; Alternative splicing;
KW   Pyrrolidone carboxylic acid.
FT   SIGNAL       1    20       BY SIMILARITY.
FT   CHAIN       21   359       CD44 ANTIGEN.
FT   DOMAIN      21   266       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   267   287       POTENTIAL.
FT   DOMAIN     288   359       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      48   119       LINK.
FT   DOMAIN     150   158       ARG/LYS-RICH (BASIC).
FT   DOMAIN     225   266       STEM.
FT   MOD_RES     21    21       PYRROLIDONE CARBOXYLIC ACID (BY
FT                              SIMILARITY).
FT   DISULFID    53   118       BY SIMILARITY.
FT   DISULFID    77    97       BY SIMILARITY.
FT   MOD_RES    289   289       PHOSPHORYLATION (BY PKC) (PARTIAL) (BY
FT                              SIMILARITY).
FT   MOD_RES    323   323       PHOSPHORYLATION (PARTIAL) (BY
FT                              SIMILARITY).
FT   CARBOHYD    25    25       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    57    57       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   100   100       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   110   110       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   120   120       N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
SQ   SEQUENCE   359 AA;  38990 MW;  BE20461C587AA34B CRC64;

  Query Match             11.2%;  Score 186;  DB 1;  Length 359;
  Best Local Similarity   24.6%;  Pred. No. 7.3e-08;
  Matches   84;  Conservative   42;  Mismatches  159;  Indels   56;  Gaps   11;

Qy         25 LRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQVETALKASFET 84
              |   :: : ::||  |:  | |     :: ||| : |:     :|    |:: ||   |||
Db         17 LSLAQIDLNITCRYAGVFHVEKNGRYSISRTEAADLCKAFNSTLPTMAQMQKALNIGFET 76

Qy         85 CSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDTWTNSCI---- 140
              |  |:: :| |||  | ||  |  |  || |     | |:  ||:|:|      |
Db         77 CRIGFI-EGHVVIPPIHPNSICAANNTGVYILTSNTS-QYDTYCFNASAPPEEDCTSVTD 134

Qy        141 -------PEIIT---------------TKDPIFNTQTATQTTEFIVSDSTYSVASPYST 177
                    |  ||              |     |  |       | |  | :  ||
Db        135 LPNAFEGPITITIVNRDGTRYTKKGEYRTNPEDINPSTPADDDVSSGSSSERSTSGGYSI 194

Qy        178 IPA--PTTTP-----------PAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENK 223
              |||  |             |    |:  |       |    |||  || ::   |
Db        195 FHTHLPTTRPTQDQSSPWVSDSPEKTPTTKDRASGGRAQTTHGSETSGHSTGSQ---EGG 251

Qy        224 AAFKNEAAGFGGVPTALLVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETK-- 281
              |:  :       :|  |::|| |    |  |  |     :          |:|:::
Db        252 ASTTSGPIRRPQIPEWLIILASL-LALALILAVC-----IAVNSRRRCGQKKKLVINNGN 305

Qy        282 -VVKEEKANDSNPNEESKKTDKNPEESKSPSKTTVRCLEAE 321
               | : ||:  |  | |:  :     :|  |:|  : : |:
Db        306 GAVDDRKASGLN-GEASRSQEMVHLVNKESSETQDQFMTAD 345


RESULT 10
PGCA_PIG
ID   PGCA_PIG       STANDARD;      PRT;   537 AA.
AC   Q29011; O18833;
DT   01-NOV-1997 (Rel. 35, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Aggrecan core protein (Cartilage-specific proteoglycan core protein)
DE   (CSPCP) (Fragments).
GN   AGC1.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE OF 1-370.
RC   TISSUE=Cartilage;
RX   MEDLINE=93038505; PubMed=1417734;
RA   Barry F.P., Gaw J.U., Young C.N., Neame P.J.;
RT   "Hyaluronan-binding region of aggrecan from pig laryngeal cartilage.
RT   Amino acid sequence, analysis of N-linked oligosaccharides and
RT   location of the keratan sulphate.";
RL   Biochem. J. 286:761-769(1992).
RN   [2]
RP   SEQUENCE OF 324-453 FROM N.A.
RC   TISSUE=Chondrocytes;
RX   MEDLINE=98209637; PubMed=9550267;
RA   Flannery C.R., Little C.B., Caterson B.;
RT   "Molecular cloning and sequence analysis of the aggrecan interglobular
RT   domain from porcine, equine, bovine and ovine cartilage: comparison of
RT   proteinase-susceptible regions and sites of keratan sulfate
RT   substitution.";
RL   Matrix Biol. 16:507-511(1998).
RN   [3]
RP   SEQUENCE OF 454-537 FROM N.A.
RC   TISSUE=Cartilage;
RX   MEDLINE=95128522; PubMed=7827755;
RA   Barry F.P., Neame P.J., Sasse J., Pearson D.;
RT   "Length variation in the keratan sulfate domain of mammalian
RT   aggrecan.";
RL   Matrix Biol. 14:323-328(1994).
CC   -!- FUNCTION: This proteoglycan is a major component of extracellular
CC       matrix of cartilagenous tissues. A major function of this protein
CC       is to resist compression in cartilage. It binds avidly to
CC       hyaluronic acid via an amino-terminal globular region. May play a
CC       regulatory role in the matrix assembly of the cartilage.
CC   -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix (By
CC       similarity).
CC   -!- DOMAIN: Two globular domains, G1 and G2, comprise the amino
CC       terminus of the proteoglycan, while another globular region, G3,
CC       makes up the COOH terminus. G1 contains link domains and thus
CC       consists of three disulfide-bonded loop structures designated as
CC       the A, B, B' motifs. G2 is similar to G1. The keratan sulfate (KS)
CC       and the chondroitin sulfate (CS) attachment domains lie between G2
CC       and G3.
CC   -!- PTM: Contains mostly chondroitin sulfate, but also keratan sulfate
CC       chains, N-linked and O-linked oligosaccharides (By similarity).
CC   -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF019757; AAC48799.1; -.
DR   EMBL; S74664; AAC60528.2; -.
DR   PIR; S78009; S78009.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000538; Link.
DR   ProDom; PD000918; Link; 2.
DR   PROSITE; PS01241; LINK; 2.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; PARTIAL.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; PARTIAL.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Glycoprotein; Proteoglycan; Repeat; Immunoglobulin domain.
FT   NON_TER       1      1
FT   DOMAIN       28    124       IG-LIKE V-TYPE DOMAIN.
FT   DOMAIN      154    231       LINK 1.
FT   DOMAIN      252    333       LINK 2.
FT   DOMAIN       32    124       G1-A.
FT   DOMAIN      136    231       G1-B.
FT   DOMAIN      237    333       G1-B'.
FT   DISULFID     35    117       BY SIMILARITY.
FT   DISULFID    159    230       BY SIMILARITY.
FT   DISULFID    183    204       BY SIMILARITY.
FT   DISULFID    257    332       BY SIMILARITY.
FT   DISULFID    281    302       BY SIMILARITY.
FT   NON_CONS    453    454
FT   DOMAIN     <454    520       KS.
FT   DOMAIN      523   >537       CS-1.
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    223    223       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    317    317       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    371    371       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    418    418       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   NON_TER     537    537
SQ   SEQUENCE   537 AA;  58708 MW;  9A4C23CCA5422F0D CRC64;

  Query Match              9.1%;  Score 150;  DB 1;  Length 537;
  Best Local Similarity   26.4%;  Pred. No. 0.0001;
  Matches   63;  Conservative   23;  Mismatches   85;  Indels   68;  Gaps   11;

Qy         50 QQLNFTEAKEACRLLGLSLAGKDQVETALKASFETCSYGWVGDGFV--VISRISPNPKCG 107
              :: . | ||   || ||  ||    |:  | :   : || ||: |  |   ||:  ||  ||
Db        246 EKFTFQEAANECRRLGARLATTGQLYLAWRGGMDMCSAGWLADRSVRYPISKARPN--CG 303

Qy        108 KNGVGVLI---------WKVPVSRQFAAYCYNSSD----------------------TWT 136
              | :||             :  | || : | ||   |                      ||
Db        304 GNLLGVRTVYLHANQTGYPDPSSR-YDAICYTGEDFVDIPENFFGVGGEEDITIQTVTWP 362

Qy        137 NSCIP-----------EIITTKDPIFN-TQTATQTTE-FIVSDSTYSVASP-------Y 175
```

```
              :  :|             :| | |:|  : || :  | |  :  | : | |
Db        363 DVELPLPRNITEGEARGTVILTVKPVFEFSPTAPEPEEPFTFAPGTGATAFFPEAENRTGE 422

Qy        176 STIP---APTTTPPAPASTSIPRRKKLICVTEVFME--------TSTMSTETEPPVENK 223
              :| |     :||   | |:      :: ||    |         |   ||  |||   |
Db        423 ATRPWAFPEESTPGLGAPTAFTSEDLVVQVTSAATEEGTEGPSATEAPSTSEEPFPSEK 481


RESULT 11
PGCA_CHICK
ID   PGCA_CHICK     STANDARD;      PRT;  2109 AA.
AC   P07898; Q90810; Q90820; Q90991; Q91047;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Aggrecan core protein precursor (Cartilage-specific proteoglycan core
DE   protein) (CSPCP).
GN   AGC1.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=White leghorn; TISSUE=Embryo;
RX   MEDLINE=94043149; PubMed=8226878;
RA   Li H., Schwartz N.B., Vertel B.M.;
RT   "cDNA cloning of chick cartilage chondroitin sulfate (aggrecan) core
RT   protein and identification of a stop codon in the aggrecan gene
RT   associated with the chondrodystrophy, nanomelia.";
RL   J. Biol. Chem. 268:23504-23511(1993).
RN   [2]
RP   SEQUENCE OF 1042-1559 FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=90307744; PubMed=1694853;
RA   Krueger R.C. Jr., Fields T.A., Mensch J.R. Jr., Schwartz N.B.;
RT   "Chick cartilage chondroitin sulfate proteoglycan core protein. II.
RT   Nucleotide sequence of cDNA clone and localization of the S103L
RT   epitope.";
RL   J. Biol. Chem. 265:12088-12097(1990).
RN   [3]
RP   SEQUENCE OF 1-1855 AND 1893-2109 FROM N.A.
RC   TISSUE=Cartilage;
RX   MEDLINE=93111968; PubMed=1339285;
RA   Chandrasekaran L., Tanzer M.L.;
RT   "Molecular cloning of chicken aggrecan. Structural analyses.";
RL   Biochem. J. 288:903-910(1992).
RN   [4]
RP   ERRATUM.
RX   MEDLINE=94107258; PubMed=8280087;
RA   Chandrasekaran L., Tanzer M.L.;
RL   Biochem. J. 296:885-887(1993).
RN   [5]
RP   SEQUENCE OF 1492-1610 FROM N.A.
RC   STRAIN=White leghorn; TISSUE=Chondrocytes;
RX   MEDLINE=95128519; PubMed=7827752;
RA   Primorac D., Stover M.L., Clark S.H., Rowe D.W.;
RT   "Molecular basis of nanomelia, a heritable chondrodystrophy of
RT   chicken.";
RL   Matrix Biol. 14:297-305(1994).
RN   [6]
RP   SEQUENCE OF 1894-2109 FROM N.A.
RX   MEDLINE=89008500; PubMed=3170613;
RA   Tanaka T., Har-El R., Tanzer M.L.;
RT   "Partial structure of the gene for chicken cartilage proteoglycan
RT   core protein.";
RL   J. Biol. Chem. 263:15831-15835(1988).
RN   [7]
RP   SEQUENCE OF 1693-1855 AND 1893-2109 FROM N.A.
RX   MEDLINE=86259736; PubMed=3460082;
RA   Sai S., Tanaka T., Kosher R.A., Tanzer M.L.;
RT   "Cloning and sequence analysis of a partial cDNA for chicken
RT   cartilage proteoglycan core protein.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:5081-5085(1986).
CC   -!- FUNCTION: This proteoglycan is a major component of extracellular
CC       matrix of cartilagenous tissues. A major function of this protein
CC       is to resist compression in cartilage. It binds avidly to
CC       hyaluronic acid via an amino-terminal globular region. May play a
CC       regulatory role in the matrix assembly of the cartilage.
CC   -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix (By
CC       similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P07898-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=P07898-2; Sequence=VSP_003073;
CC   -!- DOMAIN: Two globular domains, G1 and G2, comprise the amino
CC       terminus of the proteoglycan, while another globular region, G3,
CC       makes up the COOH terminus. G1 contains link domains and thus
CC       consists of three disulfide-bonded loop structures designated as
CC       the A, B, B' motifs. G2 is similar to G1. The keratan sulfate (KS)
CC       and the chondroitin sulfate (CS) attachment domains lie between G2
CC       and G3.
CC   -!- PTM: Contains mostly chondroitin sulfate, but also keratan sulfate
CC       chains, N-linked and O-linked oligosaccharides.
CC   -!- DISEASE: DEFECTS IN AGC1 ARE THE CAUSE OF NANOMELIA, A LETHAL
CC       CONNECTIVE TISSUE DISORDER AFFECTING CARTILAGE DEVELOPMENT
CC       (CHONDRODYSTROPHY) CHARACTERIZED BY SHORTENED AND MALFORMED LIMBS.
CC       AGGRECAN IS TRUNCATED AT ITS C-TERMINAL IN THE CS-2 BINDING DOMAIN
CC       AND IS NOT ANYMORE SECRETED FROM THE CHONDROCYTES.
CC   -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC   -!- SIMILARITY: Contains 4 link domains.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L21913; AAB19128.1; -.
DR   EMBL; M38187; AAA48731.1; -.
DR   EMBL; M88101; -; NOT_ANNOTATED_CDS.
DR   EMBL; S74657; AAC60751.1; -.
DR   EMBL; S74656; AAC60751.1; JOINED.
DR   EMBL; J04028; AAA48719.1; -.
DR   EMBL; M13993; AAA48720.1; -.
DR   PIR; I50421; I50421.
DR   HSSP; P08709; 1BF9.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000538; Link.
DR   InterPro; IPR003324; SGXXSG.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF02339; SGXXSG; 56.
DR   Pfam; PF00084; sushi; 1.
DR   Pfam; PF00193; Xlink; 4.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 4.
```

```
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00445; LINK; 4.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS50026; EGF_3; 1.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS01241; LINK; 4.
KW   Glycoprotein; Proteoglycan; Lectin; Signal; Sushi; EGF-like domain;
KW   Alternative splicing; Repeat; Immunoglobulin domain.
FT   SIGNAL        1     16       POTENTIAL.
FT   CHAIN        17   2109       AGGRECAN CORE PROTEIN.
FT   DOMAIN       34    143       IG-LIKE V-TYPE.
FT   DOMAIN      166    243       LINK 1.
FT   DOMAIN      264    346       LINK 2.
FT   DOMAIN      537    614       LINK 3.
FT   DOMAIN      635    716       LINK 4.
FT   DOMAIN     1363   1742       19 X 20 AA TANDEM-REPEAT.
FT   DOMAIN     1855   1892       EGF-LIKE.
FT   DOMAIN     1901   2019       C-TYPE LECTIN.
FT   DOMAIN     2023   2081       SUSHI.
FT   DOMAIN       48    137       G1-A.
FT   DOMAIN      148    243       G1-B.
FT   DOMAIN      249    346       G1-B'.
FT   DOMAIN      519    613       G2-B.
FT   DOMAIN      620    715       G2-B'.
FT   DOMAIN      718    803       KS.
FT   DOMAIN      805   1264       CS-1.
FT   DOMAIN     1265   1742       CS-2.
FT   DOMAIN     1893   2109       G3.
FT   DISULFID     51    129       BY SIMILARITY.
FT   DISULFID    171    242       BY SIMILARITY.
FT   DISULFID    195    216       BY SIMILARITY.
FT   DISULFID    269    345       BY SIMILARITY.
FT   DISULFID    293    314       BY SIMILARITY.
FT   DISULFID    542    613       BY SIMILARITY.
FT   DISULFID    566    587       BY SIMILARITY.
FT   DISULFID    640    715       BY SIMILARITY.
FT   DISULFID    664    685       BY SIMILARITY.
FT   DISULFID   1859   1870       BY SIMILARITY.
FT   DISULFID   1864   1879       BY SIMILARITY.
FT   DISULFID   1881   1890       BY SIMILARITY.
FT   DISULFID   1897   1908       BY SIMILARITY.
FT   DISULFID   1925   2017       BY SIMILARITY.
FT   DISULFID   1993   2009       BY SIMILARITY.
FT   DISULFID   2024   2067       BY SIMILARITY.
FT   DISULFID   2053   2080       BY SIMILARITY.
FT   CARBOHYD     76     76       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    122    122       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    330    330       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    388    388       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    439    439       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    644    644       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    700    700       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    765    765       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    801    801       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC   1856   1892       Missing (in isoform 2).
FT                                /FTId=VSP_003073.
FT   CONFLICT    362    362       E -> D (IN REF. 3).

  Query Match             9.0%;  Score 148.5;  DB 1;  Length 2109;
  Best Local Similarity   27.0%;  Pred. No. 0.00075;
  Matches   74;  Conservative   27;  Mismatches  104;  Indels   69;  Gaps   13;

Qy         50 QQLNFTEAKEACRLLGLSLAGKDQVETALKASFETCSYGWVGDGFV--VISRISPNPKCG 107
              ::  | || : |  ||  ||   ::  | |   : || ||: |  |   |||  ||  ||
Db        258 EKFTFQEAFDKCHSLGARLATTGELYLAWKDGMDMCSAGWLADRSVRYPISRARPN--CG 315

Qy        108 KNGVGV-LIWKVPV--------SRQFAAYCYNSSDTWTNSCIP-----EIITTKDPIFNT 153
              | |||  ::  |          | :: | ||:  |    : :|      |: |      |
Db        316 GNLVGVRTVYLNPANQTGYPHPSSRYDAICYSGDD--FEALVPGLFTDEVGTELGSAFTI 373

Qy        154 QTATQT----------TEFIVSDSTYSVASPYSTIPAPTTTPPAPASTSIPRRKKLICVT 203
              || |||           ||   ::  |:|:        | |    | | :|     :  |
Db        374 QTVTQTEVELPLPRNVTE---EEARGSIATLEPMEITATATELYEAFTVLPD----LFAT 426

Qy        204 EVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTALLVLALLFFGAAAGLGFCYVKRYV 263
              | :||::       |  ||    : |  |   ||                      | |
Db        427 SVTVETAS------PREEN--VTREEITGIWAVPEE------------------VTTSV 459

Qy        264 KAFPFTNKNQQKEMIETKVVKEEKANDSNPNEES 297
                ||       | |   |:|  |  : |  ||
Db        460 SGTAFTT-----GMAEVSSVEEAIAVTATPGLES 488


RESULT 12
PGCV_CHICK
ID   PGCV_CHICK     STANDARD;      PRT;  3562 AA.
AC   Q90953; Q90945;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Versican core protein precursor (Large fibroblast proteoglycan)
DE   (Chondroitin sulfate proteoglycan core protein 2) (PG-M).
GN   CPSG2.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS V0 AND V1).
RC   STRAIN=White leghorn; TISSUE=Limb bud;
RX   MEDLINE=93300846; PubMed=8314802;
RA   Shinomura T., Nishida Y., Ito K., Kimata K.;
RT   "cDNA cloning of PG-M, a large chondroitin sulfate proteoglycan
RT   expressed during chondrogenesis in chick limb buds. Alternative
RT   spliced multiforms of PG-M and their relationships to versican.";
RL   J. Biol. Chem. 268:14461-14469(1993).
CC   -!- FUNCTION: May play a role in intercellular signaling and in
CC       connecting cells with the extracellular matrix. May take part in
CC       the regulation of cell motility, growth and differentiation. Binds
CC       hyaluronic acid.
CC   -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC         Comment=Additional isoforms seem to exist;
CC       Name=V0;
CC         IsoId=Q90953-1; Sequence=Displayed;
CC       Name=V1;
CC         IsoId=Q90953-2; Sequence=VSP_003093;
CC   -!- TISSUE SPECIFICITY: Prechondrogenic condensation area of
CC       developing limb buds.
CC   -!- DEVELOPMENTAL STAGE: Disappears after the cartilage development
CC       (By similarity).
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC   -!- SIMILARITY: Contains 2 link domains.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X60226; CAA42787.1; -.
DR   EMBL; D13542; BAA02742.1; -.
```

```
DR   PIR; A47171; A47171.
DR   HSSP; P00740; 1EDM.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR000538; Link.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF00084; sushi; 1.
DR   Pfam; PF00193; Xlink; 2.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 2.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00445; LINK; 2.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 2.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 2.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS01241; LINK; 2.
KW   Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;
KW   Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
KW   Hyaluronic acid; Alternative splicing.
FT   SIGNAL        1     26       POTENTIAL.
FT   CHAIN        27   3562       VERSICAN CORE PROTEIN.
FT   DOMAIN       27    143       IG-LIKE V-TYPE.
FT   DOMAIN      166    243       LINK 1.
FT   DOMAIN      264    345       LINK 2.
FT   DOMAIN     3254   3290       EGF-LIKE 1.
FT   DOMAIN     3292   3328       EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     3341   3455       C-TYPE LECTIN.
FT   DOMAIN     3460   3518       SUSHI.
FT   DISULFID     44    129       BY SIMILARITY.
FT   DISULFID    171    242       BY SIMILARITY.
FT   DISULFID    195    216       BY SIMILARITY.
FT   DISULFID    269    344       BY SIMILARITY.
FT   DISULFID    293    314       BY SIMILARITY.
FT   DISULFID   3258   3269       BY SIMILARITY.
FT   DISULFID   3263   3278       BY SIMILARITY.
FT   DISULFID   3280   3289       BY SIMILARITY.
FT   DISULFID   3296   3307       BY SIMILARITY.
FT   DISULFID   3301   3316       BY SIMILARITY.
FT   DISULFID   3318   3327       BY SIMILARITY.
FT   DISULFID   3334   3345       BY SIMILARITY.
FT   DISULFID   3362   3454       BY SIMILARITY.
FT   DISULFID   3430   3446       BY SIMILARITY.
FT   DISULFID   3461   3504       BY SIMILARITY.
FT   DISULFID   3490   3517       BY SIMILARITY.
FT   CARBOHYD    163    163       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    235    235       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    329    329       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    529    529       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    709    709       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    948    948       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1409   1409       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1479   1479       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1523   1523       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1530   1530       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1625   1625       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1751   1751       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1988   1988       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2088   2088       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2089   2089       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2507   2507       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2642   2642       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2679   2679       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2748   2748       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2762   2762       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3069   3069       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3194   3194       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3232   3232       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3545   3545       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    485   1411       Missing (in isoform V1).
FT                                /FTId=VSP_003093.
SQ   SEQUENCE   3562 AA;  388078 MW;  9BC566E88C1602D2 CRC64;

  Query Match             8.8%;  Score 145.5;  DB 1;  Length 3562;
  Best Local Similarity   26.1%;  Pred. No. 0.0025;
  Matches   47;  Conservative   25;  Mismatches   89;  Indels   19;  Gaps    5;

Qy         50 QQLNFTEAKEACRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKN 109
              ::| | |||| ||      ||    :  | :  |: | |||: || |        |:||
Db        258 EKLTFEEAKELCRKRDGVLASVGNMYVAWRNGFDQCDYGWLADGSVRYPASVARPQCGGG 317

Qy        110 GVGVLI---------WKVPVSRQFAAYCYNSSDTWTNSCIPEIITT--KDPIFNTQTATQ 158
               :||           :  | |: | ||||      :     :::||   | :  :
Db        318 LLGVRTLYRYENQTGFPYPDSK-FDAYCYERKKIVSEPTTVKLVTTLKTDSVELSSAKVT 376

Qy        159 TTEFIVSDSTYSVASPYSTIP----APTTTPPAPASTSIPRRKKLICVTEVFMETSTMST 214
                :   |   ||   : :|    |   |    :| :   |: :   || ||   ::|
Db        377 LKPSVFESSVTEVAVTKTKVPAWEEATLETEDTKMTTEVAEEKREM---EVLMENIKLTT 433


RESULT 13
TSG6_HUMAN
ID   TSG6_HUMAN     STANDARD;      PRT;   277 AA.
AC   P98066; Q8WWI9;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Tumor necrosis factor-inducible protein TSG-6 precursor (TNF-
DE   stimulated gene 6 protein) (Hyaluronate-binding protein).
GN   TNFAIP6 OR TSG6.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fibroblast;
RX   MEDLINE=92112993; PubMed=1730767;
RA   Lee T.H., Wisniewski H.-G., Vilcek J.;
RT   "A novel secretory tumor necrosis factor-inducible protein (TSG-6) is
RT   a member of the family of hyaluronate binding proteins, closely
RT   related to the adhesion receptor CD44.";
RL   J. Cell Biol. 116:545-557(1992).
RN   [2]
RP   SEQUENCE FROM N.A., AND VARIANT GLN-144.
RX   MEDLINE=21975206; PubMed=11854277;
RA   Nentwich H.A., Mustafa Z., Rugg M.S., Marsden B.D., Cordell M.R.,
RA   Mahoney D.J., Jenkins S.C., Dowling B., Fries E., Milner C.M.,
RA   Loughlin J., Day A.J.;
RT   "A novel allelic variant of the human TSG-6 gene encoding an amino
RT   acid difference in the CUB module. Chromosomal localization,
RT   frequency analysis, modeling, and expression.";
RL   J. Biol. Chem. 277:15354-15362(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung, and Spleen;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
```

```
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [4]
RP   STRUCTURE BY NMR OF 36-133.
RX   MEDLINE=96390850; PubMed=8797823;
RA   Kohda D., Morton C.J., Parkar A.A., Hatanaka H., Inagaki F.M.,
RA   Campbell I.D., Day A.J.;
RT   "Solution structure of the link module: a hyaluronan-binding domain
RT   involved in extracellular matrix stability and cell migration.";
RL   Cell 86:767-775(1996).
CC   -!- FUNCTION: Possibly involved in cell-cell and cell-matrix
CC       interactions during inflammation and tumorgenesis.
CC   -!- INDUCTION: By TNF.
CC   -!- SIMILARITY: Contains 1 link domain.
CC   -!- SIMILARITY: Contains 1 CUB domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M31165; AAB00792.1; -.
DR   EMBL; AJ421518; CAD13434.1; -.
DR   EMBL; AJ419936; CAD12353.1; -.
DR   EMBL; BC030205; AAH30205.1; -.
DR   PIR; A41735; A41735.
DR   PDB; 1O7B; 07-NOV-03.
DR   Genew; HGNC:11898; TNFAIP6.
DR   MIM; 600410; -.
DR   GO; GO:0004895; F:cell adhesion receptor activity; TAS.
DR   GO; GO:0005540; F:hyaluronic acid binding; TAS.
DR   GO; GO:0007267; P:cell-cell signaling; TAS.
DR   GO; GO:0006954; P:inflammatory response; TAS.
DR   GO; GO:0007165; P:signal transduction; TAS.
DR   InterPro; IPR000859; CUB.
DR   InterPro; IPR000538; Link.
DR   Pfam; PF00431; CUB; 1.
DR   Pfam; PF00193; Xlink; 1.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 1.
DR   SMART; SM00042; CUB; 1.
DR   SMART; SM00445; LINK; 1.
DR   PROSITE; PS01180; CUB; 1.
DR   PROSITE; PS01241; LINK; 1.
KW   Cell adhesion; Signal; Glycoprotein; Polymorphism; 3D-structure.
FT   SIGNAL       1    19       POTENTIAL.
FT   CHAIN       20   277       TUMOR NECROSIS FACTOR-INDUCIBLE PROTEIN
FT                              TSG-6.
FT   DOMAIN      53   128       LINK.
FT   DOMAIN     135   247       CUB.
FT   DISULFID    58   127
FT   DISULFID    82   103
FT   DISULFID   135   161       BY SIMILARITY.
FT   DISULFID   188   210       BY SIMILARITY.
FT   CARBOHYD   118   118       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   258   258       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT    144   144       R -> Q.
FT                              /FTId=VAR_013005.
FT   STRAND      38    39
FT   HELIX       51    60
FT   TURN        71    72
FT   STRAND      85    85
FT   STRAND      88    88
FT   TURN        89    90
FT   STRAND      91    91
FT   TURN       103   106
FT   TURN       116   117
FT   STRAND     124   124
FT   STRAND     126   127
SQ   SEQUENCE   277 AA;  31231 MW;  4DB3AEB4AC52B880 CRC64;

  Query Match             8.8%;  Score 145;  DB 1;  Length 277;
  Best Local Similarity   33.0%;  Pred. No. 0.00012;
  Matches   29;  Conservative   16;  Mismatches   41;  Indels    2;  Gaps    1;

Qy         45 SKKANQQLNFTEAKEACRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNP 104
              ::    :| : ||| |   |   ||   |:| | |  |  |: ||:  | |    : | |
Db         42 ARSGKYKLTYAEAKAVCEFEGGHLATYKQLEAARKIGFHVCAAGWMAKGRVGYPIVKPGP 101

Qy        105 KCGKNGVGVLIWKVPVSR--QFAAYCYN 130
                ||    |:: : : ::|  :: |||||
Db        102 NCGFGKTGIIDYGIRLNRSERWDAYCYN 129


RESULT 14
TSG6_MOUSE
ID   TSG6_MOUSE     STANDARD;      PRT;   275 AA.
AC   O08859;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Tumor necrosis factor-inducible protein TSG-6 precursor (TNF-
DE   stimulated gene 6 protein).
GN   TNFAIP6 OR TNFIP6 OR TSG6.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND DEVELOPMENTAL STAGE.
RC   STRAIN=CD-1; TISSUE=Cumulus cell, Embryo, and Oocyte;
RX   MEDLINE=98087423; PubMed=9427551;
RA   Fueloep C., Kamath R.V., Li Y., Otto J.M., Salustri A., Olsen B.R.,
RA   Glant T.T., Hascall V.C.;
RT   "Coding sequence, exon-intron structure and chromosomal localization
RT   of murine TNF-stimulated gene 6 that is specifically expressed by
RT   expanding cumulus cell-oocyte complexes.";
RL   Gene 202:95-102(1997).
CC   -!- FUNCTION: Possibly involved in cell-cell and cell-matrix
CC       interactions during inflammation and tumorgenesis (By
CC       similarity).
CC   -!- DEVELOPMENTAL STAGE: Expressed in cumulus cell-oocyte complexes
CC       during expansion in vivo.
CC   -!- SIMILARITY: Contains 1 link domain.
CC   -!- SIMILARITY: Contains 1 CUB domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U83903; AAC53527.1; -.
DR   PIR; JC6506; JC6506.
DR   HSSP; P98066; 1TSG.
DR   MGD; MGI:1195266; Tnfaip6.
DR   InterPro; IPR000859; CUB.
```

```
DR   InterPro; IPR000538; Link.
DR   Pfam; PF00431; CUB; 1.
DR   Pfam; PF00193; Xlink; 1.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 1.
DR   SMART; SM00042; CUB; 1.
DR   SMART; SM00445; LINK; 1.
DR   PROSITE; PS01180; CUB; 1.
DR   PROSITE; PS01241; LINK; 1.
KW   Cell adhesion; Signal; Glycoprotein.
FT   SIGNAL        1    17       POTENTIAL.
FT   CHAIN        18   275       TUMOR NECROSIS FACTOR-INDUCIBLE PROTEIN
FT                               TSG-6.
FT   DOMAIN       35   128       LINK.
FT   DOMAIN      135   247       CUB.
FT   DISULFID     58   127       BY SIMILARITY.
FT   DISULFID     82   103       BY SIMILARITY.
FT   DISULFID    135   161       BY SIMILARITY.
FT   DISULFID    188   210       BY SIMILARITY.
FT   CARBOHYD    118   118       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    258   258       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   275 AA;  30924 MW;  1CD247228260B8F9 CRC64;

  Query Match             8.7%;  Score 144;  DB 1;  Length 275;
  Best Local Similarity   33.0%;  Pred. No. 0.00014;
  Matches   29;  Conservative   16;  Mismatches   41;  Indels    2;  Gaps    1;

Qy         45 SKKANQQLNFTEAKEACRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNP 104
              ::    :| : |||  |   |  ||  :|:| | |  |  |: ||:  | |    : | |
Db         42 ARAGRYKLTYAEAKAVCEFEGGRLATYKQLEAARKIGFHVCAAGWMAKGRVGYPIVKPGP 101

Qy        105 KCGKNGVGVLIWKVPVSR--QFAAYCYN 130
                ||    |:: : : ::|  :: |||||
Db        102 NCGFGKTGIIDYGIRLNRSERWDAYCYN 129


RESULT 15
PGCA_BOVIN
ID   PGCA_BOVIN     STANDARD;      PRT;  2364 AA.
AC   P13608; P79117; Q28159;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Aggrecan core protein precursor (Cartilage-specific proteoglycan core
DE   protein) (CSPCP).
GN   AGC1.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Hering T.M., Kollar J., Huynh T.D.;
RL   Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 563-1056 FROM N.A.
RX   MEDLINE=89380219; PubMed=2528543;
RA   Antonsson P., Heinegaard D., Oldberg A.;
RT   "The keratan sulfate-enriched region of bovine cartilage proteoglycan
RT   consists of a consecutively repeated hexapeptide motif.";
RL   J. Biol. Chem. 264:16170-16173(1989).
RN   [3]
RP   SEQUENCE OF 1609-2113 AND 2151-2364 FROM N.A.
RX   MEDLINE=87270630; PubMed=3111460;
RA   Oldberg A., Antonsson P., Heinegaard D.;
RT   "The partial amino acid sequence of bovine cartilage proteoglycan,
RT   deduced from a cDNA clone, contains numerous Ser-Gly sequences
RT   arranged in homologous repeats.";
RL   Biochem. J. 243:255-259(1987).
RN   [4]
RP   SEQUENCE OF 2114-2150 FROM N.A.
RC   TISSUE=Cartilage;
RX   MEDLINE=93352525; PubMed=8349621;
RA   Fueloep C., Walcz E., Valyon M., Glant T.T.;
RT   "Expression of alternatively spliced epidermal growth factor-like
RT   domains in aggrecans of different species. Evidence for a novel
RT   module.";
RL   J. Biol. Chem. 268:17377-17383(1993).
RN   [5]
RP   PARTIAL SEQUENCE.
RX   MEDLINE=85027710; PubMed=6489519;
RA   Perin J.-P., Bonnet F., Jolles J., Jolles P.;
RT   "Sequence data concerning the protein core of the cartilage
RT   proteoglycan monomers. Characterization of a sequence allowing the
RT   synthesis of an oligonucleotide probe.";
RL   FEBS Lett. 176:37-42(1984).
RN   [6]
RP   PARTIAL SEQUENCE.
RX   MEDLINE=87005253; PubMed=3530809;
RA   Perin J.P., Bonnet F., Jolles P.;
RT   "Structural relationship between link proteins and proteoglycan
RT   monomers.";
RL   FEBS Lett. 206:73-77(1986).
CC   -!- FUNCTION: This proteoglycan is a major component of extracellular
CC       matrix of cartilagenous tissues. A major function of this protein
CC       is to resist compression in cartilage. It binds avidly to
CC       hyaluronic acid via an amino-terminal globular region. May play a
CC       regulatory role in the matrix assembly of the cartilage.
CC   -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix (By
CC       similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P13608-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=P13608-2; Sequence=VSP_003072;
CC   -!- DOMAIN: Two globular domains, G1 and G2, comprise the amino
CC       terminus of the proteoglycan, while another globular region, G3,
CC       makes up the C-terminus. G1 contains link domains and thus
CC       consists of three disulfide-bonded loop structures designated as
CC       the A, B, B' motifs. G2 is similar to G1. The keratan sulfate (KS)
CC       and the chondroitin sulfate (CS) attachment domains lie between G2
CC       and G3.
CC   -!- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO N-LINKED AND O-
CC       LINKED (ABOUT 40) OLIGOSACCHARIDES.
CC   -!- PTM: THE KERATAN SULFATE CONTENTS DIFFER CONSIDERABLY BETWEEN
CC       ADULT AND FETAL BOVINE PROTEOGLYCANS.
CC   -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC   -!- SIMILARITY: Contains 4 link domains.
CC   -!- SIMILARITY: Contains 1 EGF-like domain.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC   -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U76615; AAB38524.1; -.
DR   EMBL; L07053; -; NOT_ANNOTATED_CDS.
DR   PIR; A34234; A39808.
DR   PIR; T42630; T42630.
DR   HSSP; P08709; 1BF9.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001304; Lectin_C.
```

```
DR   InterPro; IPR000538; Link.
DR   InterPro; IPR003324; SGXXSG.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00059; lectin_c; 1.
DR   Pfam; PF02339; SGXXSG; 61.
DR   Pfam; PF00084; sushi; 1.
DR   Pfam; PF00193; Xlink; 4.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 4.
DR   SMART; SM00032; CCP; 1.
DR   SMART; SM00034; CLECT; 1.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00445; LINK; 4.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS50026; EGF_3; 1.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; FALSE_NEG.
DR   PROSITE; PS01241; LINK; 4.
KW   Glycoprotein; Proteoglycan; Lectin; Signal; Sushi; EGF-like domain;
KW   Calcium; Alternative splicing; Repeat; Immunoglobulin domain.
FT   SIGNAL        1     16       POTENTIAL.
FT   CHAIN        17   2364       AGGRECAN CORE PROTEIN.
FT   DOMAIN       25    147       IG-LIKE V-TYPE.
FT   DOMAIN      170    247       LINK 1.
FT   DOMAIN      268    349       LINK 2.
FT   DOMAIN      504    581       LINK 3.
FT   DOMAIN      602    683       LINK 4.
FT   DOMAIN      774    907       23 X 6 AA APPROXIMATE TANDEM REPEATS OF
FT                                E-[EK]-P-F-P-S.
FT   DOMAIN     1433   2112       CS-2.
FT   DOMAIN     2113   2149       EGF-LIKE, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     2114   2364       G3.
FT   DOMAIN     2161   2276       C-TYPE LECTIN.
FT   DOMAIN     2280   2338       SUSHI.
FT   DISULFID     51    133       BY SIMILARITY.
FT   DISULFID    175    246       BY SIMILARITY.
FT   DISULFID    199    220       BY SIMILARITY.
FT   DISULFID    273    348       BY SIMILARITY.
FT   DISULFID    297    318       BY SIMILARITY.
FT   DISULFID    509    580       BY SIMILARITY.
FT   DISULFID    533    554       BY SIMILARITY.
FT   DISULFID    607    682       BY SIMILARITY.
FT   DISULFID    631    652       BY SIMILARITY.
FT   DISULFID   2117   2128       BY SIMILARITY.
FT   DISULFID   2182   2274       BY SIMILARITY.
FT   DISULFID   2250   2266       BY SIMILARITY.
FT   DISULFID   2281   2324       BY SIMILARITY.
FT   DISULFID   2310   2337       BY SIMILARITY.
FT   CARBOHYD    126    126       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    239    239       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    333    333       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    387    387       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    611    611       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    667    667       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC   2114   2150       Missing (in isoform 2).
FT                                /FTId=VSP_003072.
SQ   SEQUENCE   2364 AA;  246359 MW;  6FF83763420C3D4C CRC64;

  Query Match             8.6%;  Score 143;  DB 1;  Length 2364;
  Best Local Similarity   29.0%;  Pred. No. 0.0024;
  Matches   67;  Conservative   18;  Mismatches  100;  Indels   46;  Gaps   11;

Qy         50 QQLNFTEAKEACRLLGLSLAGKDQVETALKASFETCSYGWVGDGFV--VISRISPNPKCG 107
              ::  | ||    || ||  ||    |:  | :   : || ||: |  |    ||:  ||  ||
Db        262 EKFTFQEAANECRRLGARLATTGQLYLAWQGGMDMCSAGWLADRSVRYPISKARPN--CG 319

Qy        108 KNGVGVLI---------WKVPVSRQFAAYCYNSSDTWTNSCIPE-----------IITT 146
              | :||              :  | || : | ||    |      |||         | |
Db        320 GNLLGVRTVYLHANQTGYPDPSSR-YDAICYTGEDFVD---IPESFFGVGGEEDITIQTV 375

Qy        147 KDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTPPAPASTSIPRRKKLICVTEVF 206
               |          ||      |    | |   : :||    |     | :|  :        |
Db        376 TWPDVELPLPRNITEGEARGSVILTAKPDFEV-SPTAPEPEEPFTFVPEVR-----ATAF 429

Qy        207 METSTMSTE-TEPFVENKAAFKNEAAGFGGVPTA-----LLVLALLFFGAA 251
               |    : | | |:     ||  |:    | |||     |:|   |  |||
Db        430 PEVENRTEEATRPW-----AFPRESTPGLGAPTAFTSEDLVVQVTLAPGAA 475


Search completed: August 11, 2004, 11:53:18
Job time : 14 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 11, 2004, 11:49:26 ; Search time 39 Seconds
(without alignments)
2605.048 Million cell updates/sec

Title: US-10-063-510-6
Perfect score: 1657
Sequence: 1 MARCFSLVLLLTSIWTTRLL.....NPEESKSPSKTTVRCLEAEV 322

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1657	100.0	322	4	Q9UNF4	Q9unf4 homo sapien
2	1643	99.2	322	4	Q8TC18	Q8tc18 homo sapien
3	1638	98.9	322	4	Q9Y5Y7	Q9y5y7 homo sapien
4	1103	66.6	318	11	Q8BHC0	Q8bhc0 mus musculu
5	1095	66.1	318	11	Q99NE4	Q99ne4 mus musculu
6	726.5	43.8	201	6	Q7YS22	Q7ys22 sus scrofa
7	227	13.7	364	11	O70509	O70509 rattus norv
8	219.5	13.2	780	11	O08779	O08779 rattus norv
9	219	13.2	265	13	Q98SR5	Q98sr5 anas platyr
10	217.5	13.1	398	13	Q90ZL8	Q90zl8 anas platyr
11	206.5	12.5	168	13	Q90ZL6	Q90zl6 anas platyr
12	204	12.3	361	4	Q86T72	Q86t72 homo sapien
13	203.5	12.3	364	6	O97569	O97569 ceratotheri
14	203.5	12.3	396	13	Q9W6S4	Q9w6s4 gallus gall
15	203	12.3	742	4	Q9UJ36	Q9uj36 homo sapien
16	202	12.2	361	4	Q8N694	Q8n694 homo sapien
17	201	12.1	580	11	Q80X37	Q80x37 mus musculu
18	199.5	12.0	493	4	Q86Z27	Q86z27 homo sapien
19	199	12.0	338	4	Q9H5A4	Q9h5a4 homo sapien
20	197	11.9	719	4	Q9H5A5	Q9h5a5 homo sapien
21	196.5	11.9	699	4	Q96J24	Q96j24 homo sapien
22	196	11.8	294	4	Q92493	Q92493 homo sapien
23	191.5	11.6	676	4	Q9H5A7	Q9h5a7 homo sapien
24	191	11.5	271	4	Q9H5A3	Q9h5a3 homo sapien
25	191	11.5	470	4	Q9H5A6	Q9h5a6 homo sapien
26	149.5	9.0	2109	13	P79787	P79787 gallus gall
27	145.5	8.8	494	6	Q9BGH3	Q9bgh3 sus scrofa
28	145	8.8	1069	4	Q9UF98	Q9uf98 homo sapien
29	140.5	8.5	1238	11	Q8K0K6	Q8k0k6 mus musculu
30	140.5	8.5	2571	11	Q8R4Y4	Q8r4y4 mus musculu
31	136	8.2	894	11	Q8BM87	Q8bm87 mus musculu
32	136	8.2	2559	11	Q8R4U0	Q8r4u0 mus musculu
33	134	8.1	897	4	Q9NRY3	Q9nry3 homo sapien
34	133	8.0	514	11	Q62913	Q62913 rattus norv
35	133	8.0	1192	4	Q9H7H7	Q9h7h7 homo sapien
36	133	8.0	1416	4	Q86UR4	Q86ur4 homo sapien
37	133	8.0	1431	11	Q8CFM6	Q8cfm6 rattus norv
38	133	8.0	1736	4	Q8TES1	Q8tes1 homo sapien
39	133	8.0	2551	4	Q8WWQ8	Q8wwq8 homo sapien
40	133	8.0	2551	4	Q8IUG9	Q8iug9 homo sapien
41	133	8.0	2551	4	Q7Z5N9	Q7z5n9 homo sapien
42	131	7.9	911	4	Q96FP7	Q96fp7 homo sapien
43	131	7.9	911	4	Q96GW7	Q96gw7 homo sapien
44	130.5	7.9	883	11	Q80WT7	Q80wt7 mus musculu
45	130.5	7.9	1152	13	Q90WM2	Q90wm2 xenopus lae

ALIGNMENTS

RESULT 1
Q9UNF4
ID Q9UNF4 PRELIMINARY; PRT; 322 AA.
AC Q9UNF4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hyaluronic acid receptor.
GN HAR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Winkelmann J.C., Basu S., Ozdemir E., Blough R.I.;
RT "HAR: a novel homolog of CD44 and putative hyaluronic acid receptor
RT encoded by a gene on human chromosome 11p15.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF127670; AAD49220.2; -.
DR HSSP; P98066; 1TSG.
DR GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR000538; Link.
DR Pfam; PF00193; Xlink; 1.
DR PRINTS; PR01265; LINKMODULE.
DR ProDom; PD000918; Link; 1.
DR SMART; SM00445; LINK; 1.
KW Receptor.
SQ SEQUENCE 322 AA; 35213 MW; 8B4D6D623F52D559 CRC64;

Query Match 100.0%; Score 1657; DB 4; Length 322;
Best Local Similarity 100.0%; Pred. No. 9.8e-142;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
```

```
Qy          61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy         121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy         181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy         241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy         301 DKNPEESKSPSKTTVRCLEAEV 322
               |||||||||||||||||||||
Db         301 DKNPEESKSPSKTTVRCLEAEV 322


RESULT 2
Q8TC18
ID   Q8TC18      PRELIMINARY;      PRT;   322 AA.
AC   Q8TC18;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Extracellular link domain-containing 1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Strausberg R.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC026231; AAH26231.1; -.
DR   Genew; HGNC:14687; XLKD1.
DR   GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   InterPro; IPR000538; Link.
DR   Pfam; PF00193; Xlink; 1.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 1.
DR   SMART; SM00445; LINK; 1.
SQ   SEQUENCE   322 AA;  35183 MW;  0B1EDBD76CE4610A CRC64;

  Query Match             99.2%;  Score 1643;  DB 4;  Length 322;
  Best Local Similarity   99.7%;  Pred. No. 1.8e-140;
  Matches  321;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy           1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy          61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db          61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIRKVPV 120

Qy         121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy         181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy         241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy         301 DKNPEESKSPSKTTVRCLEAEV 322
               |||||||||||||||||||||
Db         301 DKNPEESKSPSKTTVRCLEAEV 322


RESULT 3
Q9Y5Y7
ID   Q9Y5Y7      PRELIMINARY;      PRT;   322 AA.
AC   Q9Y5Y7;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Lymphatic endothelium-specific hyaluronan receptor LYVE-1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99156989; PubMed=10037799;
RA   Banerji S., Ni J., Wang S.X., Clasper S., Su J., Tammi R., Jones M.,
RA   Jackson D.G.;
RT   "LYVE-1, a new homologue of the CD44 glycoprotein, is a lymph-specific
RT   receptor for hyaluronan.";
RL   J. Cell Biol. 144:789-801(1999).
DR   EMBL; AF118108; AAD42764.1; -.
DR   HSSP; P98066; 1TSG.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0005624; C:membrane fraction; TAS.
DR   GO; GO:0004872; F:receptor activity; TAS.
DR   GO; GO:0006928; P:cell motility; TAS.
DR   GO; GO:0007160; P:cell-matrix adhesion; TAS.
DR   GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.
DR   GO; GO:0009611; P:response to wounding; TAS.
DR   InterPro; IPR000538; Link.
DR   Pfam; PF00193; Xlink; 1.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 1.
DR   SMART; SM00445; LINK; 1.
KW   Receptor.
SQ   SEQUENCE   322 AA;  35238 MW;  0EBEA56729CEFFF7 CRC64;

  Query Match             98.9%;  Score 1638;  DB 4;  Length 322;
  Best Local Similarity   99.1%;  Pred. No. 5.1e-140;
  Matches  319;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy           1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy          61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
               |||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db          61 CRLLGLSLAGKDQVDTALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy         121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
               |||| |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db         121 SRQFGAYCYNSSDTWTNSRIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy         181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy         241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy         301 DKNPEESKSPSKTTVRCLEAEV 322
               |||||||||||||||||||||
Db         301 DKNPEESKSPSKTTVRCLEAEV 322


RESULT 4
```

```
Q8BHC0
ID   Q8BHC0      PRELIMINARY;      PRT;   318 AA.
AC   Q8BHC0;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Extra cellular link domain-containing 1.
GN   XLKD1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CZECH II; TISSUE=Lung, and Mammary gland;
RA   Strausberg R.;
RL   Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Lung;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; BC038653; AAH38653.1; -.
DR   EMBL; BC038892; AAH38892.1; -.
DR   EMBL; AK004726; BAC25094.1; -.
DR   MGD; MGI:2136348; Xlkd1.
DR   GO; GO:0005886; C:plasma membrane; IDA.
DR   GO; GO:0005540; F:hyaluronic acid binding; IDA.
DR   GO; GO:0004888; F:transmembrane receptor activity; IDA.
DR   GO; GO:0006027; P:glycosaminoglycan catabolism; IDA.
DR   InterPro; IPR000538; Link.
DR   Pfam; PF00193; Xlink; 1.
DR   ProDom; PD000918; Link; 1.
DR   SMART; SM00445; LINK; 1.
DR   PROSITE; PS01241; LINK; 1.
SQ   SEQUENCE   318 AA;  34573 MW;  34AA31AEF5430B08 CRC64;

  Query Match           66.6%;  Score 1103;  DB 11;  Length 318;
  Best Local Similarity   69.7%;  Pred. No. 1.5e-91;
  Matches  221;  Conservative   30;  Mismatches   62;  Indels    4;  Gaps    3;

Qy          6 SLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLG 65
              |||||| ||||||  |||:   ::|||  :|||||: || :  | |:||||| |||::||
Db          6 SLVLLLASIWTTRHPVQGADLVQDLSIS-TCRIMGVALVGRNKNPQMNFTEANEACKMLG 64

Qy         66 LSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFA 125
              |:|| :||||:| |: ||||||||||: | || ||  ||:||||| |||||  | |::|
Db         65 LTLASRDQVESAQKSGFETCSYGWVGEQFSVIPRIFSNPRCGKNGKGVLIWNAPSSQKFK 124

Qy        126 AYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTP 185
              |||:|||||| |||||||:||  |: :||  |  ||| || | |  :|| || |   ||
Db        125 AYCHNSSDTWVNSCIPEIVTTFYPVLDTQ--TPATEFSVSSSAYLASSPDSTTPVSATT- 181

Qy        186 PAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTALLVLAL 245
              ||  ||: |: | ||:|||: |  ||:|||| || : |||||||||||||||||||||
Db        182 RAPPLTSMARKTKKICITEVYTEPITMATETEAFVASGAAFKNEAAGFGGVPTALLVLAL 241

Qy        246 LFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKTDKNPE 305
              ||||||| |  |||||||||||| ||||||||||||||||:| | ||||||| ||||
Db        242 LFFGAAAVLAVCYVKRYVKAFPFTTKNQQKEMIETKVVKEEKADDVNANEESKKTIKNPE 301

Qy        306 ESKSPSKTTVRCLEAEV 322
              |:||| ||||||||||
Db        302 EAKSPPKTTVRCLEAEV 318


RESULT 5
Q99NE4
ID   Q99NE4      PRELIMINARY;      PRT;   318 AA.
AC   Q99NE4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hyaluronan receptor precursor.
GN   XLKD1 OR LYVE-1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Balb/c; TISSUE=Digestive tract;
RX   MEDLINE=99156989; PubMed=10037799;
RA   Banerji S.;
RT   "LYVE-1, a new homologue of the CD44 glycoprotein is a lymph-specific
RT   receptor for hyaluronan.";
RL   J. Cell Biol. 144:789-801(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Balb/c; TISSUE=Digestive tract;
RX   MEDLINE=21276443; PubMed=11278811;
RA   Prevo R., Banerji S., Ferguson D.J.P., Clasper S., Jackson D.G.;
RT   "Mouse LYVE-1 is an endocytic receptor for hyaluronan in lymphatic
RT   endothelium.";
RL   J. Biol. Chem. 276:19420-19430(2001).
DR   EMBL; AJ311501; CAC33082.1; -.
DR   HSSP; P98066; 1TSG.
DR   MGD; MGI:2136348; Xlkd1.
DR   GO; GO:0005886; C:plasma membrane; IDA.
DR   GO; GO:0005540; F:hyaluronic acid binding; IDA.
DR   GO; GO:0004888; F:transmembrane receptor activity; IDA.
DR   GO; GO:0006027; P:glycosaminoglycan catabolism; IDA.
DR   InterPro; IPR000538; Link.
DR   Pfam; PF00193; Xlink; 1.
DR   ProDom; PD000918; Link; 1.
DR   SMART; SM00445; LINK; 1.
DR   PROSITE; PS01241; LINK; 1.
KW   Receptor; Signal.
FT   SIGNAL       1     23
FT   CHAIN       24    318       LYVE-1.
SQ   SEQUENCE   318 AA;  34641 MW;  1248974A16113330 CRC64;

  Query Match           66.1%;  Score 1095;  DB 11;  Length 318;
  Best Local Similarity   69.1%;  Pred. No. 7.8e-91;
  Matches  219;  Conservative   30;  Mismatches   64;  Indels    4;  Gaps    3;

Qy          6 SLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLG 65
              |||  | ||||||  |||:   ::|||  :|||||: || :  | |:||||| |||::||
Db          6 SLVFFLASIWTTRHPVQGADLVQDLSIS-TCRIMGVALVGRNKNPQMNFTEANEACKMLG 64

Qy         66 LSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFA 125
              |:|| :||||:| |: ||||||||||: | || ||  ||:||||| |||||  | |::|
Db         65 LTLASRDQVESAQKSGFETCSYGWVGEQFSVIPRIFSNPRCGKNGKGVLIWNAPSSQKFK 124

Qy        126 AYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTP 185
              |||:|||||| |||||||:||  |: :||  |  ||| || | |  :|| || |   ||
Db        125 AYCHNSSDTWVNSCIPEIVTTFYPVLDTQ--TPATEFSVSSSAYLASSPDSTTPVSATT- 181

Qy        186 PAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTALLVLAL 245
              ||  ||: |: | ||:|||: |  ||:|||| || : |||||||||||||||||||||
Db        182 RAPPLTSMARKTKKICITEVYTEPITMATETEAFVASGAAFKNEAAGFGGVPTALLVLAL 241

Qy        246 LFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKTDKNPE 305
              ||||||| |  |||||||||||| ||||||||||||||||:| | ||||||| ||||
Db        242 LFFGAAAVLAVCYVKRYVKAFPFTTKNQQKEMIETKVVKEEKADDVNANEESKKTIKNPE 301

Qy        306 ESKSPSKTTVRCLEAEV 322
              |:||| ||||||||||
Db        302 EAKSPPKTTVRCLEAEV 318


RESULT 6
```

```
Q7YS22
ID   Q7YS22      PRELIMINARY;      PRT;   201 AA.
AC   Q7YS22;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Lymphatic endothelial hyaluronan receptor LYVE-1 (Fragment).
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Michael K.W., Xu S.-H., Voisine P., Khan T.A., Feng J., Li J.,
RA   Sellke F.W., Bianchi C.;
RT   "Identification of pig LYVE-1.";
RL   Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY304537; AAP69946.1; -.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER     201    201
SQ   SEQUENCE   201 AA;  22219 MW;  BAAF34E7EACBBAE3 CRC64;

  Query Match             43.8%;  Score 726.5;  DB 6;  Length 201;
  Best Local Similarity   70.1%;  Pred. No. 1e-57;
  Matches  141;  Conservative   21;  Mismatches   36;  Indels    3;  Gaps    2;

Qy          81 SFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDTWTNSCI 140
               |||||||||| | |:|| || ||||||||| |||:|:  :|::| |||:||||| |||||
Db           1 SFETCSYGWVKDQFLVIPRIYPNPKCGKNGKGVLVWRHSLSQKFEAYCHNSSDTRTNSCI 60

Qy         141 PEIITTKDPIFNTQTATQTTEFIVSDSTYSVAS--PYSTI-PAPTTTPPAPASTSIPRRK 197
               |||| | || ||| ||  |||  |:| | | ::  | |:: |    ||     |:|| ||::
Db          61 PEIIPTNDPTFNTNTAPYTTEMTVNDRTSSSSTNGPSSSVMPTVVTTSLPLATTSTPRKR 120

Qy         198 KLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTALLVLALLFFGAAAGLGFC 257
               ||||:|| |||||:||||| ::||: |||||| ||||:||||||||||| |||||   |
Db         121 KLICITEAFMETSTISTETELYIENRTAFKNEAIGFGGIPTALLVLALLFFAAAAGLAVC 180

Qy         258 YVKRYVKAFPFTNKNQQKEMI 278
               |||||||||||||||||||||
Db         181 YVKRYVKAFPFTNKNQQKEMI 201


RESULT 7
O70509
ID   O70509      PRELIMINARY;      PRT;   364 AA.
AC   O70509;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Glycoprotein CD44s.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Lumbar spine;
RA   Stevens J.W.;
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF065147; AAC17117.1; -.
DR   HSSP; P98066; 1TSG.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR   GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   InterPro; IPR001231; CD44_antigen.
DR   InterPro; IPR000538; Link.
DR   Pfam; PF00193; Xlink; 1.
DR   PRINTS; PR00658; CD44.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 1.
DR   SMART; SM00445; LINK; 1.
DR   PROSITE; PS01241; LINK; 1.
SQ   SEQUENCE   364 AA;  39725 MW;  BA249776C4419AA7 CRC64;

  Query Match             13.7%;  Score 227;  DB 11;  Length 364;
  Best Local Similarity   24.3%;  Pred. No. 4e-12;
  Matches   86;  Conservative   55;  Mismatches  153;  Indels   60;  Gaps   13;

Qy          15 WTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQV 74
               |    |:| ||  ::: : ::||  |:  | |      :: ||| : |     :|    |:
Db          10 WGLLCLLQLSLAQQQIDLNITCRYAGVFHVEKNGRYSISRTEAADLCEAFNTTLPTMAQM 69

Qy          75 ETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDT 134
               | ||   |||| ||:: :| ||| || ||  |  |  || |      :  :  ||:|:|
Db          70 ELALSKGFETCRYGFI-EGHVVIPRIHPNAICAANNTGVYILLASNTSHYDTYCFNASAP 128

Qy         135 WTNSCIPEIITTKDPIFNTQTATQTTEFIVSDST-YSVASPYST----IPAPT------- 182
                    |     |:   : |:     |   :  | | ||   | |     | |
Db         129 LEEDC-----TSVTDLPNSFDGPVTITIVNRDGTRYSKKGEYRTHQEDIDASNIIDEDVS 183

Qy         183 -------TTPPA-PASTSIPRRKKLICVTEVFMETSTMST-ETEPFVENKAAFK-----N 228
                    :||      | :|  :      : |   ||::| : :  :: :  |      :
Db         184 SGSTIEKSTPEGYILHTDLPTSQPTGDRDDAFFIGSTLATSDGDSSMDPRGGFDTVTHGS 243

Qy         229 EAAGFGG-----------------VPTALLVLALLFFGAAAGLGFCYVKRYVKAFPFTN 270
               | ||                  :|  |::|| |     |  | |  |     :
Db         244 ELAGHSSGNQDSGVTTTSGPARRPQIPEWLIILASL-LALALILAVC-----IAVNSRRR 297

Qy         271 KNQQKEMIETK---VVKEEKANDSNPNEESKKTDKNPEESKSPSKTTVRCLEAE 321
                 |:|:::      |:: | :: |  | || :    :| |::|  : : |:
Db         298 CGQKKKLVINSGNGTVEDRKPSELN-GEASKSQEMVHLVNKEPTETPDQFMTAD 350


RESULT 8
O08779
ID   O08779      PRELIMINARY;      PRT;   780 AA.
AC   O08779;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   CD44 protein.
GN   CD44.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BDX; TISSUE=Pancreas;
RA   Hofmann M.;
RT   "Rattus norvegicus CD44 protein sequence.";
RL   Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U96138; AAB54002.1; -.
DR   HSSP; P98066; 1TSG.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR   GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   InterPro; IPR001231; CD44_antigen.
DR   InterPro; IPR000538; Link.
DR   Pfam; PF00193; Xlink; 1.
DR   PRINTS; PR00658; CD44.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 1.
DR   SMART; SM00445; LINK; 1.
DR   PROSITE; PS01241; LINK; 1.
SQ   SEQUENCE   780 AA;  85917 MW;  CC4D35AB1EA7377C CRC64;

  Query Match             13.2%;  Score 219.5;  DB 11;  Length 780;
  Best Local Similarity   24.9%;  Pred. No. 5.1e-11;
  Matches   86;  Conservative   45;  Mismatches  144;  Indels   71;  Gaps   13;

Qy          15 WTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQV 74
```

```
Db         10 WGLLCLLQLSLAQQQIDLNITCRYAGVFHVEKNGRYSISRTEAADLCEAFNTTLPTMAQM 69

Qy         75 ETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDT 134

Db         70 ELALRKGFETCRYGFI-EGHVVIPRIHPNAICAANNTGVYILLASNTSHYDTYCFNASAP 128

Qy        135 WTNSCIPEIITTKDPIFNTQTATQTTEFIVSDST-YSVASPYST----IPAPT------- 182

Db        129 LEEDC-----TSVTDLPNSFDGPVTITIVNRDGTRYSKKGEYRTHQEDIDASNIIDEDVS 183

Qy        183 -------TTPPA-PASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAG-- 232

Db        184 SGSTIEKSTPEGYILHTDLPTSQPTGDRDDAFFIGSTLAT-----IASTVYSKSHATAQK 238

Qy        233 --------FGG---------VPTALLVLALLFFGAA------AGLGFCYVKRYVKAFPFT 269

Db        239 QNNWIWSWFGNSQSTTQTQDSPTTTATTALMTTPETPPKRQEAQNWFSW---------FF 289

Qy        270 NKNQQKEMIETKVVKEEKANDSN-----PNEESK-KTDKNPEESKS 309

Db        290 QPSESKSHLHTTTKMPGTESNTNPTGWKPNEENEDETDKYPNFSGS 335
```

RESULT 9
Q98SR5
ID Q98SR5 PRELIMINARY; PRT; 265 AA.
AC Q98SR5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE T cell antigen CD44 isoform b.
OS Anas platyrhynchos (Domestic duck).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anas.
OX NCBI_TaxID=8839;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White Pekin;
RA Chan S.W.S., Warr G.W., Middleton D.L., Higgins D.A.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF332869; AAK18277.1; -.
DR HSSP; P98066; 1TSG.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR001231; CD44_antigen.
DR InterPro; IPR000538; Link.
DR Pfam; PF00193; Xlink; 1.
DR PRINTS; PR00658; CD44.
DR PRINTS; PR01265; LINKMODULE.
DR ProDom; PD000918; Link; 1.
DR SMART; SM00445; LINK; 1.
DR PROSITE; PS01241; LINK; 1.
SQ SEQUENCE 265 AA; 29215 MW; 77C176E0A898D081 CRC64;

Query Match 13.2%; Score 219; DB 13; Length 265;
Best Local Similarity 28.3%; Pred. No. 1.4e-11;
Matches 66; Conservative 32; Mismatches 103; Indels 32; Gaps 5;

```
Qy         14 IWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQ 73

Db          6 VWATFGLCLLKLCLTETQFNVSCRYRGVFHVEKNGRYSLTRTEAADLCRALNSTLSTLEQ 65

Qy         74 VETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSD 133

Db         66 LEKAHELGFETCRYGFV-VGYIVIPRINPYHLCAANHTGIYKLSANTTGRYDAYCYNATE 124

Qy        134 TWTNSCIP------------EIITTKDP-----IFNTQTATQTTEFIVSDSTYSVASPY 175

Db        125 TRDKACEPIERIDTSFLSNQGEIVIDNEDGSRYNADGTRHSGDSSTSGVDDENVGSGSSH 184
```

```
Qy        176 STIPAPTTTPPAPAS-----TSIPR--------RKKLICVTEVFMETSTMSTE 215

Db        185 DTTPVDTSIRRSSPSYYGSVTPVPHLSDHSSGGGEKEFPVTNSDDEISPTSTD 237
```

RESULT 10
Q90ZL8
ID Q90ZL8 PRELIMINARY; PRT; 398 AA.
AC Q90ZL8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE T cell antigen CD44 isoform a.
OS Anas platyrhynchos (Domestic duck).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anas.
OX NCBI_TaxID=8839;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White Pekin;
RA Chan S.W.S., Middleton D.L., Warr G.W., Higgins D.A.;
RT "Anas platyrhynchos T cell antigens.";
RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY029553; AAK40246.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR001231; CD44_antigen.
DR InterPro; IPR000538; Link.
DR Pfam; PF00193; Xlink; 1.
DR PRINTS; PR00658; CD44.
DR PRINTS; PR01265; LINKMODULE.
DR ProDom; PD000918; Link; 1.
DR SMART; SM00445; LINK; 1.
DR PROSITE; PS01241; LINK; 1.
SQ SEQUENCE 398 AA; 43673 MW; 25A944EE2F4AED6D CRC64;

Query Match 13.1%; Score 217.5; DB 13; Length 398;
Best Local Similarity 29.8%; Pred. No. 3.2e-11;
Matches 56; Conservative 28; Mismatches 85; Indels 19; Gaps 3;

```
Qy         14 IWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQ 73

Db          6 VWATFGLCLLKLCLTETQFNVSCRYRGVFHVEKNGRYSLTRTEAADLCRALNSTLSTLEQ 65

Qy         74 VETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSD 133

Db         66 LEKAHELGFETCRYGFV-VGYIVIPRINPYHLCAANHTGIYKLSANTTGRYDAYCYNATE 124

Qy        134 TWTNSCIP------------EIITTKDP-----IFNTQTATQTTEFIVSDSTYSVASPY 175

Db        125 TRDKACEPIERIDTSFLSNQGEIVIDNEDGSRYNADGTRHSGDSSTSGVDDENVGSGSSH 184

Qy        176 STIPAPTT 183

Db        185 DTTPVDTS 192
```

RESULT 11
Q90ZL6
ID Q90ZL6 PRELIMINARY; PRT; 168 AA.
AC Q90ZL6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE T-cell antigen CD44 isoform c.
OS Anas platyrhynchos (Domestic duck).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anas.
OX NCBI_TaxID=8839;
RN [1]
RP SEQUENCE FROM N.A.

```
RA   Chan S.W.S., Middleton D.L., Lundqvist M., Warr G.W., Higgins D.A.;
RT   "Anas platyrhynchos T-cell antigens.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY032667; AAK52086.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR   GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   InterPro; IPR001231; CD44_antigen.
DR   InterPro; IPR000538; Link.
DR   Pfam; PF00193; Xlink; 1.
DR   PRINTS; PR00658; CD44.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 1.
DR   SMART; SM00445; LINK; 1.
DR   PROSITE; PS01241; LINK; 1.
SQ   SEQUENCE   168 AA;  18964 MW;  122EFE727977056E CRC64;

  Query Match             12.5%;  Score 206.5;  DB 13;  Length 168;
  Best Local Similarity   35.2%;  Pred. No. 1.1e-10;
  Matches   45;  Conservative   22;  Mismatches   60;  Indels    1;  Gaps    1;

Qy          14 IWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQ 73
               :| |  |     |   :    |||:  |:  | |      |  ||| : || |  :|:  :|
Db           6 VWATFGLCLLKLCLTKTQFNVSCKYRGVFHVEKNGRYSLTRTEAADLCRALNSTLSTLEQ 65

Qy          74 VETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSD 133
               :| | :  |||| ||:|  |::|| ||:|   |  |  |:          : :: |||||::::
Db          66 LEKAHELGFETCRYGFV-VGYIVIPRINPYHLCAANHTGIYKLSANTTGRYDAYCYNATE 124

Qy         134 TWTNSCIP 141
               |   :| |
Db         125 TRDKACEP 132


RESULT 12
Q86T72
ID   Q86T72      PRELIMINARY;      PRT;   361 AA.
AC   Q86T72;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein DKFZp451K1918.
GN   DKFZP451K1918.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ansorge W., Krieger S., Mewes H.W., Weil B., Amid C., Fobo G., Han M.,
RA   Osanger A., Wiemann S.;
RL   Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL832642; CAD89965.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR   GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   InterPro; IPR001231; CD44_antigen.
DR   InterPro; IPR000538; Link.
DR   Pfam; PF00193; Xlink; 1.
DR   PRINTS; PR00658; CD44.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 1.
DR   SMART; SM00445; LINK; 1.
DR   PROSITE; PS01241; LINK; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   361 AA;  39387 MW;  F51A746B44387C54 CRC64;

  Query Match             12.3%;  Score 204;  DB 4;  Length 361;
  Best Local Similarity   24.4%;  Pred. No. 4.8e-10;
  Matches   83;  Conservative   46;  Mismatches  145;  Indels   66;  Gaps   14;

Qy          25 LRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQVETALKASFET 84
               |   :: : ::||  |:  | |     :: ||| : |:    :|     |:| ||   |||
Db          17 LSLAQIDLNITCRFAGVFHVEKNGRYSISRTEAADLCKAFNSTLPTMAQMEKALSIGFET 76

Qy          85 CSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDTWTNSCIPEII 144
               | ||:: :| ||| || ||  |  |  || |     | |:  ||:|:|      |
Db          77 CRYGFI-EGHVVIPRIHPNSICAANNTGVYILTSNTS-QYDTYCFNASAPPEEDC----- 129

Qy         145 TTKDPIFNTQTATQTTEFIVSDST-YSVASPYSTIPA---PTTTPPAPASTSIPRRKKLI 200
               |:   : |       |   :  | | |     | | |    |:       |:     :
Db         130 TSVTDLPNAFDGPITITIVNRDGTRYVQKGEYRTNPEDIYPSNPTDDDVSSGSSSERSST 189

Qy         201 CVTEVFMETSTM-------------STETEPFVENKAAFK----------NEAAGF---- 233
                  :|   ||:             ||:  |   ::  |           :|: |
Db         190 SGGYIFYTFSTVHPIPDEDSPWITDSTDRIPATRDQDTFHPSGGSHTTHGSESDGHSHGS 249

Qy         234 --GG------------VPTALLVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIE 279
                 ||           :|  |::|| |    |  |  |     :          |:|:::
Db         250 QEGGANTTSGPIRTPQIPEWLIILASL-LALALILAVC-----IAVNSRRRCGQKKKLVI 303

Qy         280 TK---VVKEEKANDSNPNEESKKTD----KNPEESKSPSK 312
               |:: | :  |  | ||  :     | | |::| :
Db         304 NSGNGAVEDRKPSGLN-GEASKSQEMVHLANKESSETPDQ 342


RESULT 13
O97569
ID   O97569      PRELIMINARY;      PRT;   364 AA.
AC   O97569;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   CD44 (Fragment).
OS   Ceratotherium simum simum.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Ceratotherium.
OX   NCBI_TaxID=73337;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Worley M.B.;
RT   "White rhinoceros CD44.";
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF045939; AAD02418.1; -.
DR   HSSP; P98066; 1TSG.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR   GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   InterPro; IPR001231; CD44_antigen.
DR   InterPro; IPR000538; Link.
DR   Pfam; PF00193; Xlink; 1.
DR   PRINTS; PR00658; CD44.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 1.
DR   SMART; SM00445; LINK; 1.
DR   PROSITE; PS01241; LINK; 1.
FT   NON_TER     364     364
SQ   SEQUENCE   364 AA;  39670 MW;  C56B5E410AAB1CFF CRC64;

  Query Match             12.3%;  Score 203.5;  DB 6;  Length 364;
  Best Local Similarity   25.4%;  Pred. No. 5.4e-10;
  Matches   91;  Conservative   49;  Mismatches  141;  Indels   77;  Gaps   17;

Qy          20 LVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQVETALK 79
               |||  ||    :: : ::||  |:  | |      :: ||| : |:    :|    |:| ||
Db          14 LVQLSL--AQIDLNITCRYAGVFHVEKNGRYSISRTEAADLCQAFNSTLPTMAQMEKALS 71

Qy          80 ASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSS------D 133
                  |||| ||:: :| ||| || ||  |  |  || |     | |:  ||:|:|      |
Db          72 IGFETCRYGFI-EGHVVIPRIHPNSICAANNTGVYILTSNTS-QYDTYCFNASAPPGEED 129

Qy         134 TWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDST-YSVASPYSTIPA---PTTTPPAPA 189
                 : : :|    : ||        |   :  | | ||    | | |    |:|
```

```
Db        130 CSSVTDLPNAF--EGPI--------TITIVNRDGTRYSKKGEYRTNPEDINPSTQADEDV 179

Qy        190 STSIPRRKKLICVTEVF----METSTMSTETEPFV----ENKAAFKN------------ 228
              |:     :       :|      |     :  |:|    ||    |:
Db        180 SSGSSSERSTSGGYSIFHTHLPTTRPTQDQGSPWVSDHSENTPTTKDQGSSDDHSGRSHT 239

Qy        229 ----EAAGF------GG-----------VPTALLVLALLFFGAAAGLGFCYVKRYVKAF 266
                  |: |:      ||           :|  |::|| |    |  |  |     :
Db        240 THGTESPGYSRGSQEGGANTTSGPIRKPQIPEWLIILASL-LALALILAVC-----IAVN 293

Qy        267 PFTNKNQQKEMIETK---VVKEEKANDSNPNEESKKTDKNPEESKSPSKTTVRCLEAE 321
                    |:|:::       |:: ||:  |  | ||  :     :|  |:|  : : |:
Db        294 SRRRCGQKKKLVINNGNGAVEDRKASGLN-GEASKSQEMVHLVNKESSETQDQFMTAD 350


RESULT 14
Q9W6S4
ID   Q9W6S4      PRELIMINARY;      PRT;   396 AA.
AC   Q9W6S4;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   CD44-like protein.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Amnion;
RA   Zhu S., Dong S., Halfter W.;
RT   "Transient expression of a CD44-like protein in the optic nerve and
RT   chiasm of the chick embryo.";
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF153205; AAD37443.1; -.
DR   HSSP; P98066; 1TSG.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR   GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   InterPro; IPR001231; CD44_antigen.
DR   InterPro; IPR000538; Link.
DR   Pfam; PF00193; Xlink; 1.
DR   PRINTS; PR00658; CD44.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 1.
DR   SMART; SM00445; LINK; 1.
DR   PROSITE; PS01241; LINK; 1.
SQ   SEQUENCE   396 AA;  43217 MW;  652CD73322605516 CRC64;

  Query Match             12.3%;  Score 203.5;  DB 13;  Length 396;
  Best Local Similarity   22.9%;  Pred. No. 6e-10;
  Matches   82;  Conservative   42;  Mismatches  135;  Indels   99;  Gaps   11;

Qy         29 ELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQVETALKASFETCSYG 88
              |    ::||  |:  | |      |   ||| | || |  :||  :| | |    |||| ||
Db         21 ETQFNITCRYGGVFHVEKNGRYSLTRAEAIELCRALNSTLATLEQFERAHALGFETCRYG 80

Qy         89 WVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDTWTNSCIP------- 141
              ::  | :|| ||:|    |  |  |:       : :: |||||:::| : :| |
Db         81 FI-VGHIVIPRINPYHLCAANHTGIYKLSANTTGRYDAYCYNATETRSKACEPIERIDIT 139

Qy        142 ------EIIT----------------------TKDPIFNTQTATQTTEFIV-----SDS 167
                   ||:                      |    : :   ||        | |
Db        140 FLSNQSEIVIDNEDGSRYNADGTRHSGDSSTSGVDDENLGSGSIHDTTPGDASIRRSSPS 199

Qy        168 TYSVASPYSTIP-------------------APTTTPPAPASTSIPRRKKLICVTEVFM 207
              |   :||| :|                  :||:|    :   ||         :|
Db        200 YYGSVTPYSHMPDHSSGGGEKDFPVKHYDDEISPTSTDILATAADFPRE------DDVQH 253

Qy        208 ETSTMSTE---------------TEPFVENKAAFKNEAAGFGGVPTALLVLALLFFGAA 251
                 || ||               : | :       ::  |   ||  |:::| |    |

Db        254 PASTRSTSNIDSDQGPHKGDGEPTSSPGLSTTTTVTSQ-PGTAHVPEWLIIVAAL-LPLA 311

Qy        252 AGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKTDKNPEESKS 309
                 | |     :          |:|:::          |     |: |  : | : |||
Db        312 LILAVC-----IAVNSRRRCGQKKKLVIN--------NGKGAVEDRKTRELNGDASKS 356


RESULT 15
Q9UJ36
ID   Q9UJ36      PRELIMINARY;      PRT;   742 AA.
AC   Q9UJ36;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Transmembrane glycoprotein precursor.
GN   CD44.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gunthert U.;
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94147793; PubMed=7508842;
RA   Gunthert U.;
RT   "CD44: a multitude of isoforms with diverse functions.";
RL   Curr. Top. Microbiol. Immunol. 184:47-63(1993).
DR   EMBL; AJ251595; CAB61878.1; -.
DR   PIR; A47195; A47195.
DR   HSSP; P98066; 1TSG.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR   GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   InterPro; IPR001231; CD44_antigen.
DR   InterPro; IPR000538; Link.
DR   Pfam; PF00193; Xlink; 1.
DR   PRINTS; PR00658; CD44.
DR   PRINTS; PR01265; LINKMODULE.
DR   ProDom; PD000918; Link; 1.
DR   SMART; SM00445; LINK; 1.
DR   PROSITE; PS01241; LINK; 1.
KW   Signal.
FT   SIGNAL        1     23       POTENTIAL.
FT   CHAIN        24    742       TRANSMEMBRANE GLYCOPROTEIN.
SQ   SEQUENCE   742 AA;  81598 MW;  C319E5CE5C0B51D3 CRC64;

  Query Match             12.3%;  Score 203;  DB 4;  Length 742;
  Best Local Similarity   26.0%;  Pred. No. 1.5e-09;
  Matches   61;  Conservative   30;  Mismatches   84;  Indels   60;  Gaps    7;

Qy         25 LRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQVETALKASFET 84
                 |   :: : ::||  |:  | |     :: ||| : |:    :|    |:| ||   |||
Db         17 LSLAQIDLNITCRFAGVFHVEKNGRYSISRTEAADLCKAFNSTLPTMAQMEKALSIGFET 76

Qy         85 CSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDTWTNSCI---- 140
              | ||:: :| ||| || ||  |  |  || |      | |:  ||:|:|      |
Db         77 CRYGFI-EGHVVIPRIHPNSICAANNTGVYILTYNTS-QYDTYCFNASAPPEEDCTSVTD 134

Qy        141 ----------------------------PEIITTKDPIFNTQTATQTTEFIVSDSTYS 170
                                          || |  :|  :  ::  ::|   :   |
Db        135 LPNAFDGPITITIVNRDGTRYVQKGEYRTNPEDIYPSNPTDDDVSSGSSSERSSTSGGY- 193

Qy        171 VASPYSTI-PAPTTTPP--------APASTSIPRRKKLICVTEVFMETSTMSTET 216
                 :   :||: | |    |         ||:|              | ||  :|||
Db        194 IFYTFSTVHPIPDEDSPWITDSTDRIPATT-------------LMSTSATATET 234


Search completed: August 11, 2004, 11:54:09
Job time : 40 secs
```

```
;  OTHER INFORMATION: Synthetic protein
US-09-902-775A-201

  Query Match           100.0%;  Score 1657;  DB 4;  Length 322;
  Best Local Similarity  100.0%;  Pred. No. 1.4e-159;
  Matches  322;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy        301 DKNPEESKSPSKTTVRCLEAEV 322
              |||||||||||||||||||||||
Db        301 DKNPEESKSPSKTTVRCLEAEV 322


RESULT 5
US-08-892-880-2
; Sequence 2, Application US/08892880
; Patent No. 5942417
;  GENERAL INFORMATION:
;    APPLICANT:  NI, JIAN
;    APPLICANT:  GENTZ, REINER L.
;    APPLICANT:  DILLON, PATRICK J.
;    TITLE OF INVENTION:  CD44-LIKE PROTEIN
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
;      STREET:  1100 NEW YORK AVENUE, NW, SUITE 600
;      CITY:  WASHINGTON
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/892,880
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/021,762
;      FILING DATE:  15-JUL-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  STEFFE, ERIC K
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1488.0490001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-371-2600
;      TELEFAX:  202-371-2540
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  322 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
```

```
;  MOLECULE TYPE:  protein
US-08-892-880-2

  Query Match           99.6%;  Score 1651;  DB 2;  Length 322;
  Best Local Similarity  99.7%;  Pred. No. 5.7e-159;
  Matches  321;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA 60

Qy         61 CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         61 CRLLGLSLAGKDQVETALKASFATCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV 120

Qy        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA 180

Qy        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL 240

Qy        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT 300

Qy        301 DKNPEESKSPSKTTVRCLEAEV 322
              |||||||||||||||||||||||
Db        301 DKNPEESKSPSKTTVRCLEAEV 322


RESULT 6
US-09-724-864-60
; Sequence 60, Application US/09724864
; Patent No. 6380362
; GENERAL INFORMATION:
;  APPLICANT:  Watson, James D
;  APPLICANT:  Murison, James G.
;  TITLE OF INVENTION:  Polynucleotides, polypeptides expressed
;  TITLE OF INVENTION:  by the polynucleotides and methods for their use.
;  FILE REFERENCE:  11000.1050U1
;  CURRENT APPLICATION NUMBER: US/09/724,864
;  CURRENT FILING DATE:  2000-11-28
;  PRIOR APPLICATION NUMBER:  U.S. No. 6380362 60/171,678
;  PRIOR FILING DATE:  1999-12-23
;  NUMBER OF SEQ ID NOS:  72
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 60
;   LENGTH: 318
;   TYPE: PRT
;   ORGANISM: Mouse
US-09-724-864-60

  Query Match           66.6%;  Score 1103;  DB 4;  Length 318;
  Best Local Similarity  69.7%;  Pred. No. 1.9e-103;
  Matches  221;  Conservative   30;  Mismatches   62;  Indels    4;  Gaps    3;

Qy          6 SLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLG 65
              |||||| ||||||  |||:   ::|||  :||||: || :  | |:|||||  |||::||
Db          6 SLVLLLASIWTTRHPVQGADLVQDLSIS-TCRIMGVALVGRNKNPQMNFTEANEACKMLG 64

Qy         66 LSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQFA 125
              |:|| :||||:| |: ||||||||||: | || ||  ||:|||||  |||||   | |::|
Db         65 LTLASRDQVESAQKSGFETCSYGWVGEQFSVIPRIFSNPRCGKNGKGVLIWNAPSSQKFK 124

Qy        126 AYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTP 185
              |||:|||||| |||||||:||  |: :||  |  ||| || | |  :|| || |   ||
Db        125 AYCHNSSDTWVNSCIPEIVTTFYPVLDTQ--TPATEPSVSSSAYLASSPDSTTPVSATT- 181

Qy        186 PAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTALLVLAL 245
              ||  ||: |: | ||:|||: |  ||:|||| || : ||||||||||||||||||||
Db        182 RAPPLTSMARKTKKICITEVYTEPITMATETEAFVASGAAFKNEAAGFGGVPTALLVLAL 241
```